

Essays on Institutions in Developing Economies

by

Xiao Yu Wang

B.A. Economics with Honors, University of Wisconsin-Madison (2008)
B.A. Mathematics with Honors, University of Wisconsin-Madison (2008)

Submitted to the Department of Economics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2013

© 2013 Xiao Yu Wang

The author hereby grants to Massachusetts Institute of Technology permission to
reproduce and
to distribute copies of this thesis document in whole or in part.

Signature of Author .



.....
Department of Economics
May 15, 2013

Certified by.....

.....
Abhijit V. Banerjee
Ford International Professor of Economics
Thesis Supervisor

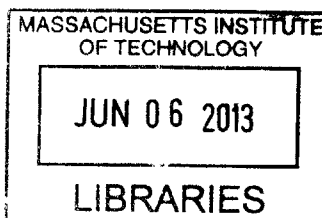
Certified by.....

.....
Robert M. Townsend
Elizabeth & James Killian Professor of Economics
Thesis Supervisor

Accepted by

.....
Michael Greenstone
3M Professor of Environmental Economics
Chairman, Departmental Committee on Graduate Studies

ARCHIVES



Essays on Institutions in Developing Economies

by

Xiao Yu Wang

Submitted to the Department of Economics
on May 15, 2013, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The primary goal of this thesis is to gain a deeper understanding of how institutional structure responds and evolves in equilibrium, particularly in the idiosyncratic and dynamic settings of developing economies. I use methods from market design to study these questions. The first chapter characterizes the equilibrium response of informal insurance relationships to changes in the formal sector, when risk-averse agents may choose what risk to bear in addition to how to share a given risk. The second chapter studies informal insurance relationships in a setting with one-sided moral hazard, and shows how the tradeoff between incentive and insurance provision shapes the composition and nature of informal relationships. The third chapter focuses on a more general setting, where the standard price mechanism fails or is not available, and provides an explanation for why the stable mechanism used in its place works well in practice, despite appearing to be easily manipulable.

In the first chapter, I develop a theory of endogenous matching between heterogeneously risk-averse individuals who, once matched, choose both the riskiness of the income stream they face (ex ante risk management) as well as how to share that risk (ex post risk management). I find a clean condition on the fundamentals of the model for unique positive-assortative and negative-assortative matching in risk attitudes. From this, I derive an intuitive falsifiability condition, discuss support for the theory in existing empirical work, and propose an experimental design to test the theory. Finally, I demonstrate the policy importance of understanding informal insurance as the risk-sharing achieved within the equilibrium network of partnerships, rather than within a single, isolated partnership. A hypothetical policy which reduces aggregate risk is a strict Pareto improvement if the matching is unchanged, but can be seen to harm the most risk-averse individuals and to exacerbate inequality when the endogenous network response is taken into account: the least risk-averse individuals abandon their roles as informal insurers in favor of entrepreneurial partnerships. This results in an increase in the risk borne by the most risk-averse agents, who must now match with each other on low-return investments.

The aim of my second chapter is to understand the impact of optimal provision of both risk and incentives on the choice of contracting partners. I study a risky setting where heterogeneously risk-averse employers and employees must match to be productive. They face a standard one-sided moral hazard problem: mean output increases in the noncontractible input of the employee. Better insurance comes at the cost of weaker incentives, and this tradeoff differs across partnerships of different risk compositions. I show that this heterogeneous tradeoff determines the equilibrium matching pattern, and focus on environments in which assortative matching is the unique equilibrium. This endogenous matching framework enables a concrete and rigorous analysis of the interaction between formal and informal insurance. In particular, I show that the introduction of formal insurance crowds out informal insurance, and may leave those individuals acting as informal

insurers in the status quo strictly worse off.

My third chapter is motivated by the observation that mechanisms which implement stable matchings often work well in practice, even in environments where individuals could gain by using simple strategies to game the mechanism. Why might individuals refrain from strategic manipulation, even when the complexity cost of manipulation is low? I study a two-sided, one-to-one matching problem with no side transfers, where utility is interdependent in the following intuitive sense: an individual's utility from a match depends not only on her preference ranking of her assigned partner, but also on that partner's ranking of her. I show that, in a world of complete information and linear interdependence, a unique stable matching emerges, and is attained by a modified Gale-Shapley deferred acceptance algorithm. As a result, a stable rule supports truth-telling as an equilibrium strategy. Hence, these results offer a new intuition for why stable matching mechanisms seem to work well in practice, despite their theoretic manipulability: individuals may value being liked.

Thesis Supervisor: Abhijit V. Banerjee

Title: Ford International Professor of Economics

Thesis Supervisor: Robert M. Townsend

Title: Elizabeth & James Killian Professor of Economics

Contents

1	Endogenous Insurance and Informal Relationships	8
1.1	Introduction	8
1.2	Background	15
1.3	The Model	17
1.4	Results	20
1.4.1	Support in the Literature	29
1.4.2	Discussion of the Model	34
1.5	Policy	36
1.5.1	Differently-sized groups	36
1.5.2	Crop price stabilization and formal insurance	37
1.6	Testing the Theory: An Experiment	43
1.6.1	Subject and Setting Choice	43
1.6.2	Setup	44
1.6.3	Empirical Strategy	50
1.7	Conclusion	53
1.A	Appendix	56
1.A.1	The model with general distribution of returns	56
1.A.2	The model nests a framework with individual project choice, pooled income, and individual rationality	59
1.A.3	Proof that NTU problem has TU representation in expected utility	59
1.A.4	Proof of main matching result	62
1.A.5	Proof that NAM is the unique equilibrium match when income-smoothing is shut down	65

1.A.6	Proof that PAM is the unique equilibrium match when consumption-smoothing is shut down	65
1.A.7	Differently-sized groups	67
1.A.8	Background and empirical context	68
2	Risk, Incentives, and Contracting Relationships	77
2.1	Introduction	77
2.2	The Model	84
2.3	Results	86
2.3.1	The First-Best	86
2.3.2	The Second-Best	91
2.4	The Policy Example	98
2.5	Conclusion	101
2.A	Appendix	104
2.A.1	The one-sided CARA-Laplace moral hazard model	104
2.A.2	The equilibrium for a given pair in the first-best	110
2.A.3	Proof of the main result in the second-best	113
3	Interdependent Utility and Truthtelling in Two-Sided Matching	115
3.1	Introduction	115
3.2	An Example	120
3.3	The Model	122
3.4	Results	125
3.5	An Example: Followup	134
3.6	Conclusion	137
	Bibliography	139

Acknowledgements

I am the fortunate beneficiary of the warmth and talent of many wonderful people. I could never have completed this dissertation without the support they so generously provided, although the social planner is probably having second thoughts about that.

If ever someone were to convince the social planner to do exactly the opposite, it would be Abhijit Banerjee. I've admired his wisdom, patience, and sense of humor since I met him the summer before my first year, when I plucked up the courage to ask if I could work as his research assistant. Although he didn't know me, and arguably had no compelling reason to get to know me, he agreed without hesitation, and immediately set me a puzzle he said had been on his mind for fifteen years. ("Of course," he added, "not continuously.") Abhijit taught me the importance of understanding the irreducible core of a problem before enveloping it in complication. He infuses seminars, ping pong matches, and dance floors with a secret delight, and he has pushed me from the beginning to be an independent researcher and to pursue the questions which interest me.

If ever the social planner had a bad day (say, discovered a sign error), and wanted to hide from the world for a while, Rob Townsend would provide shelter without judgment. He is exceptionally kind. His love of economics energizes all the people around him, and he finds interesting angles to any question. Above all, he treats everybody as an equal. Rob's commitment, his compassion, and his genuine, welcoming character enrich the entire community.

If ever the social planner had a role model, it would be Esther Duflo. That she is brilliant, highly capable, and peerless in strength of will is well-known to the point of being uninteresting, and, at the same time, an understatement. She cares deeply about her students, both as economists and as human beings. I can't express how much her support has meant to me, especially during the challenges of this past year.

In addition, I'd like to thank Ken West for guiding me as an undergraduate, when I had no idea what I was doing, and for checking up on me every year afterward, when I continued to have no idea what I was doing.

Prasanthi Durvasula and Ashley Eide are some of the best friends anyone could possibly have. I thank them for keeping me sane and laughing, and I'm really proud of them both.

For being my fourth adviser, and for writing pear jokes that make me quince, I thank Gabriel Carroll. I thank Jennifer Peck and Sarah Venables for making graduate school way more fun than it probably should have been. The stress of coursework, exams, and research was made manageable by innumerable devious campaigns, impromptu late-night meetings by the washer and dryer, and general wackiness. Sebastian Di Tella and I floundered our way through five years. I thank him for being my EBFF. He is still the only person I know who was once suspected of defrauding himself.

In graduate school, there's a lot of stuff that no one tells you, such as codes to important but frequently-locked rooms, where to find all the notes that everyone else seems to have, how to format your thesis, and so on. Somehow, you're supposed to figure everything out on your own, possibly with magic. I therefore thank Michael Powell for being magical, as well as incredibly generous. I thank Jen-Jen La'O for being the best econ big sister, and Arun Chandrasekhar for being the best econ cousin-once-removed. Dana Foarta helped me move, which is the nicest thing anyone can do for you, really. Dan Barron took me on hikes that were ex ante terrifying but ex post awesome, Marco Di Maggio fed me prosciutto smuggled from Italy, Greg Leiserson accepted me as a running buddy even though I was terrible and he was quick, and Chris Walters was my go-to guy for beer,

pizza, and bad movies. I hope to be as cool as Danielle Li and Ashley Swanson someday. Adam Sacarny, Brad Shapiro, and Juan Pablo Xandri like to give me a hard time, but are always there for me when it matters.

I've saved special thanks for a special person. In order to distract me from the aggravation of the job market, Ben Golub was kind enough to aggravate me himself. His phone calls are the sweetest, his hugs are the nicest, and his comments about the formatting of my slides are by far the most detailed. There isn't anyone else I'd rather have at my side.

Lastly, I thank my mother, Yijun, my father, Jin, and my sister, Anne, for their deep love, and for being a truly one-of-a-kind family. My parents made tremendous and incomprehensible sacrifices for me to have the opportunities denied to them. They lived through a singular period of Chinese history. I grew up listening to their childhood stories, and when I began to take economics classes, my interests naturally turned to development. For her part, my sister listens to me talk about economics, which is also a pretty big sacrifice.

Though incomparable to theirs, this dedication is for my family.

Chapter 1

Endogenous Insurance and Informal Relationships

1.1 Introduction

The distinctive features of institutions and contract structures arising in the unique environments of developing economies have inspired a vast body of research. Informal insurance alone has received a great deal of attention.¹ The poor, especially the rural poor, persistently report risk as a serious problem (Alderman and Paxson (1992), Townsend (1994), Morduch (1995), Dercon (1996), Fafchamps (2008)). There is widespread consensus that people in developing countries live in more hazardous environments (in terms of climate, disease, and landscape, for example), and thus inherently face more risk (Dercon (1996), Fafchamps (2008)). Furthermore, they live close to minimal subsistence levels. This makes them particularly vulnerable to the risks they face, since even small negative shocks could have disastrous consequences.

In addition to experiencing high levels of risk and vulnerability, the poor in developing economies often lack recourse to formal insurance and credit institutions. Consequently, they incorporate risk

¹For background and institutional details, see the classic papers of Alderman and Paxson (1992) and Morduch (1995), and the recent discussions provided by Dercon (2004) and Fafchamps (2008). Addressing the question of how well informal insurance works in practice is the seminal paper by Townsend (1994), with further papers on empirical tests of risk-sharing in a wide variety of settings by Dercon and Krishnan (2000), Fafchamps and Lund (2003), Mazzocco and Saini (2012), among many others. A further class of papers seeks theoretical explanations for the observed limitations of informal insurance. One prominent explanation is limited commitment, explored in a dynamic setting by Ligon et. al. (2002), and as a natural bound on group size by Genicot and Ray (2003), Bloch et. al. (2008). In addition, a still-growing theoretical and empirical body of literature studies risk-sharing networks—both how they form as well as how the nature of equilibrium risk-sharing corresponds to network structures (see, e.g., Bramouille and Kranton (2007), Fafchamps and Gubert (2007), Ambrus et al. (2013)).

management into their relationships with each other in a variety of creative ways. Rosenzweig and Stark (1989) show that daughters are often strategically married to households in distant villages with highly dissimilar agroclimates, so as to minimize the correlation between farming incomes. The authors show that the motivation for doing this is indeed insurance; households exposed to more income risk are more likely to invest in longer-distance marriage arrangements. Rosenzweig (1988) argues more generally that the formation of kinship networks between heterogeneously risk-averse individuals is motivated by insurance. Ligon et al. (2002) and Kocherlakota (1996), among many others, analyze a pure risk-sharing relationship between two heterogeneously risk-averse households who perfectly observe each other's income. Akerberg and Botticini (2002) study agricultural contracting in medieval Tuscany, and find evidence that heterogeneously risk-averse tenant farmers and landlords strategically formed sharecropping relationships based on differing risk attitudes, motivated by risk management concerns.

This paper enriches our understanding of informal insurance by developing and studying a theory of endogenous relationship formation between heterogeneously risk-averse individuals in risky environments, where these individuals may implement *ex ante* and *ex post* risk management strategies with each other in the absence of formal insurance and credit markets². The strength of informal insurance is thus how well-insured a population of risk-averse individuals is when they must rely only on interactions with each other to manage risk. The existing risk-sharing literature has focused largely on analyzing the insurance agreement reached by a single group of individuals, isolated outside of the equilibrium network. This provides insight into what behaviors to expect if a given group of individuals is matched, but does not provide insight into what groups could actually coexist in the first place.

Because of the lack of formal institutions and the subsequent dependence on interpersonal relationships, taking individual heterogeneity and the equilibrium network seriously is essential for our understanding of the developing world. For example, I show that accounting for a policy's induced network response, as well as its effect on the flexibility of the network to respond in the first place, may substantially alter assessment of welfare impacts. Indeed, Eeckhout and Munshi (2010) study chit funds³ in Chennai, and show that members of funds are matched strategically.

²This terminology derives from Morduch (1995), who also refers to *ex ante* risk management as "income-smoothing", and *ex post* risk management as "consumption-smoothing". Intuitively, an individual choosing the riskiness of the action she takes is managing risk *ex ante*, while an individual choosing how to smooth the riskiness of an action she has taken is managing risk *ex post*. In this paper, I use the *ex ante* and *ex post* terminology.

³Chit funds are informal financial institutions which emerge in the absence of formal credit, where groups of subscribers combine funds and bid for the pool of money.

Moreover, members swiftly re-sorted in response to an unexpected exogenous policy which imposed a cap on the bids they were allowed to place for the pot of money in their fund. The results of this paper enable policymakers to account precisely for this type of endogenous network response when designing regulations and programs.

The setting I study has the following key elements. First, a population of risk-averse individuals with CARA utility lacks access to formal insurance and credit institutions. Individuals belong to one of two groups, and members must partner up across groups in order to be productive.⁴ For example, in an agricultural village, some individuals own land but would prefer not to farm it themselves, while other, landless individuals have both the willingness and the skill to farm. No crops can be grown if landowners do not employ farmers. Alternatively, in a microentrepreneurship setting, some individuals might have skill A , while others have skill B , and a successful business venture requires the combination of both skills.

I introduce two key types of heterogeneity: heterogeneity of preferences, and heterogeneity of technology. Specifically, individuals vary in their degree of risk aversion, and can choose from an assortment of risky projects which vary in their riskiness—a riskier project has a higher expected return, but also a higher variance of return.⁵ Individuals know each other’s risk attitudes, as well as what projects are available.

Since formal institutions are absent, individuals must rely on agreements with each other to manage the risks they face. A matched pair of individuals jointly chooses one project, and commits *ex ante* to a return-contingent sharing rule. The risk composition of the pair determines both *ex ante* risk management, that is, the riskiness of the income stream the pair chooses, as well as *ex post* risk management, that is, the sharing rule describing the split of each realized return. Hence, informal insurance motivations influence the projects and contracts chosen in equilibrium, as well as the composition of partnerships ostensibly formed to produce output.

The main contributions of this paper are threefold: first, I show that this model of nontransferable utility has a transferable utility representation, and I use this representation to find conditions on the fundamentals of the modeling environment for unique assortative matching, which are *in-*

⁴In fact, I show that this model nests a model in which individuals choose their own productive opportunities, and match to pool their incomes (see Appendix 2 for the proof). However, this version of the model (with joint productivity) is more tractable and natural for a study of partner choice. Moreover, because it nests the first model, it applies to scenarios in which agents do not literally work together, such as two farmers who grow their own crops but partner up to share risk.

⁵In this model, I abstract from moral hazard, in order to focus on the impact of the heterogeneous trade-off in *ex ante* and *ex post* risk management strategies across partnerships of different risk compositions on equilibrium matching. I model moral hazard explicitly in Wang (2012b).

dependent of the distributions of risk types in the economy. Equilibrium matching is driven by the interplay between *ex ante* and *ex post* risk management strategies for a given partnership, where this interplay differs across partnerships of different risk compositions. I show that the fundamental of the model which determines this interplay is the marginal variance cost of taking up a riskier project—a riskier project has a higher mean but also a higher variance of return. The marginal variance cost describes the increase in variance incurred when moving from a project with some mean return to a project with a slightly higher mean return. A sufficient condition for unique negative-assortative matching in risk attitudes (that is, the i^{th} least risk-averse person is matched with the i^{th} most risk-averse person) is convexity of the marginal variance cost function, while concavity is sufficient for unique positive-assortative matching (that is, the i^{th} least risk-averse person is matched with the i^{th} least risk-averse person).

Because the marginal variance cost function of risky projects may be challenging to observe in the data, particularly in non-laboratory settings, I derive a simple condition to test the theory which relies only on plausibly-collectible data, such as risk attitudes (for instance, elicited using gamble choices as in Binswanger (1980) or Attanasio et al. (2012), or estimated as in Ligon et al. (2002)), network links, and some proxy of mean incomes. Specifically, I show that convexity of the marginal variance cost function is equivalent to the concavity of the mean returns of equilibrium projects in the representative risk tolerance of matched pairs. Analogously, concavity of the marginal variance cost function is equivalent to the convexity of the mean returns of equilibrium projects in the representative risk tolerance of matched pairs. Hence, given data on risk attitudes, linkages between individuals, and mean incomes, the representative risk tolerance can be calculated for each matched pair, and the behavior of mean incomes in this risk tolerance can be tested for convexity or concavity. I also propose an experimental design to test the theory directly.

Finally, I show that this theory has significant policy implications: a policy may trigger an endogenous network response, and we need a coherent and tractable model of endogeneity in order to account for this in our welfare analysis. For example, a policymaker may wish to stabilize the prices of the riskiest crops in an environment where increases in the mean returns of a crop portfolio come at an extremely steep marginal variance cost, and this trade-off is particularly onerous for crop portfolios with high mean profitability. Crop price stabilization is typically motivated by a desire to reduce the substantial risk burden of risk-averse farmers, and to encourage farmers to grow crops with higher mean profitability. This kind of policy is very common in developing countries around the world; for instance, Chile uses a variable import tariff to maintain a price band around

wheat ("2012 National Trade Estimate Report: Foreign Trade Barriers" (2011)).

I show that such a risk-reduction policy, which decreases the variance of the riskiest projects in particular, is a Pareto improvement if status quo partnerships are assumed to stay fixed. However, accounting for the endogenous network response shows that such a policy may cause particular harm to the most risk-averse agents: the least risk-averse agents abandon their roles as "informal insurers" of the most risk-averse in favor of "entrepreneurial" pursuits with other less risk-averse agents. This forces the most risk-averse agents to pair with each other, leaving them strictly worse off. Importantly, inequality is exacerbated: partnerships of the least risk-averse individuals are able to take advantage of the reduced risk and take up highly profitable projects, while the most risk-averse individuals, having only weak ability to manage risk together *ex post*, rely heavily on *ex ante* risk management, and subsequently select projects with especially low mean returns.

This work unites the literature on institutions in risky environments lacking formal insurance and credit markets with the literature on endogenous matching. A small body of literature has studied endogenous informal insurance relationships. Li et al. (2012) present a model of endogenous matching between heterogeneously risk-averse people in which both positive- and negative-assortative matching arise in equilibrium, and the certainty-equivalent is transferable within the household. However, in their model, agents bear an explicit utility cost of choosing a project with a certain mean and variance, where this utility cost function is increasing in the mean chosen and decreasing in the variance chosen. The central role of this element in the theory causes the main result to imply that the distributions of risk types drive equilibrium matching. By contrast, my results are distribution-free, and offer an alternative intuition for the key determinants of the equilibrium match. In addition, Li et al. do not discuss falsifiability of their model, or empirical applications of their results. I derive a testable condition from my main theoretical result, and show how it can be used in practice to address policy questions, and to guide empirical work in informal insurance.

Legros and Newman (2007) develop techniques to characterize stable matchings in nontransferable utility settings by generalizing the Shapley and Shubik (1972) and Becker (1974) supermodularity and submodularity conditions for matching under transferable utility. Under nontransferable utility, the indirect utility of each member of the first group given a partnership with each member of the second group can be calculated, fixing the second member's level of expected utility at some level v . Then, this indirect utility expression, which depends on both members' types and v , is analyzed for supermodularity and submodularity in risk types.

Legros and Newman also demonstrate an application of their method to a specific matching problem under nontransferable utility, that is, the problem of forming partnerships to share an exogenous risk. More specifically, heterogeneously risk-averse individuals (divided into two groups, "men" and "women") choose their partners endogenously. All matched pairs face the same exogenous risk (an individual receives either low or high income, where the probability of the low income shock is fixed and is identical for all individuals), but a matched pair is able to commit *ex ante* to a sharing rule contingent on the pooled income realization. In other words, people are able to choose how to share a given risk, but not the risk itself. They show that their generalized decreasing differences condition holds in this setting, and therefore that negative-assortative matching in risk types must be the unique equilibrium matching pattern.

Chiappori and Reny (2006) show that this result continues to hold when the exogenous risk faced by partnerships is more general than a risky income stream with binary return. More specifically, they assume that each individual is endowed with some random income stream, and that the income streams of all men are independent from and exchangeable with all income streams of women. This ensures that all women look essentially identical to all men in terms of income, and vice versa, so that the only heterogeneity driving the matching equilibrium is in risk aversion.

Schulhofer-Wohl (2006) also studies this model. He shows that this problem of endogenous matching to share an exogenous risk has a transferable utility representation iff individuals have ISHARA preferences⁶. Because the conditions for assortative matching in transferable utility settings are the standard supermodularity/submodularity conditions from Shapley and Shubik (1971), he is able to apply these standard conditions to show that negative-assortative matching is the unique equilibrium matching pattern in this model when agents have ISHARA preferences.

This paper develops and explores a model of endogenous matching between heterogeneously risk-averse individuals who are able to manage risk *ex ante* as well as *ex post*. That is, a partnership is both able to choose what risk to face, as well as how to share that risk. The direct theoretical innovation is to understand how the equilibrium structure of informal risk-sharing partnerships changes when both types of risk management strategies are available. The main result provides conditions under which negative-assortative matching arises as the unique equilibrium, *and* when positive-assortative matching arises as the unique equilibrium.

⁶To do this, he uses the Mazzocco (2007) result that the preferences of expected utility maximizers are exactly aggregable iff they are ISHARA (harmonic absolute risk aversion with identical shape parameters). (Mazzocco (2007) was in circulation as a manuscript in 2004.)

There are two important consequences of identifying an environment in which both extremal types of matching patterns can arise, and understanding the differences in the environments which lead to one matching pattern versus the other. In a recent experiment, Attanasio et al. (2012) take a rigorous experimental approach to the question of endogenous risk-sharing group formation. The key elements of the environment are limited commitment, project choice, and a sharing rule fixed at equal division. They find that individuals who are linked in kinship or friendship are more likely to form risk-sharing groups with each other than with strangers, most likely because of better quality information about risk attitudes as well as trust. Furthermore, they find that risk-sharing groups composed of individuals linked by a family or friendship tie are *positively-assorted* in risk attitude. This empirical observation of positive-assortative matching contrasts strikingly with the negative-assortative match predicted by the models of Legros and Newman, Schulhofer-Wohl, and Chiappori and Reny. Additionally, these results suggest that family and friendship ties are important for identifying the pool of potential partners for a given individual (because an individual is unlikely to know the risk attitude of, or to trust a stranger), but the choice of partner from this pool for the purpose of insurance is primarily determined by risk attitudes.

This paper provides a clear economic intuition for the unique emergence of each of the two extremal matching patterns. In addition, the falsifiability condition I derive allows empiricists to test the environment of interest for the applicability of the theory. If the model cannot be rejected, then the theory provides an understanding of the environment and sheds light on the design of the empirical study. For example, the results in this paper can be used to assess the validity of the assumption of homogeneous risk attitudes in a household⁷, which is fairly standard in work on empirical tests of risk-sharing, where data is often at the household level.

Furthermore, this framework enables the kinds of policy analysis discussed earlier (risk-reduction policy such as price stabilization, or the introduction of informal insurance), because it demonstrates precisely how the network of relationships in the status quo may shift from one composition to another, in response to changes in the environment. Hence, this paper contributes an understanding of the features of the environment which are essential for equilibrium relationship formation, in addition to the network response.

The rest of the paper proceeds as follows. In the next section, I describe the institution of sharecropping in Aurepalle, to fix the ideas of the model in a real world setting. Following this,

⁷Homogeneity of risk attitudes in a matched pair is equivalent to positive-assortative matching when the distribution of risk types in each group is the same.

I set up the model, state and discuss the main results, provide support for the theory in existing literature, and discuss key elements of the model. I then show that a policy which decreases aggregate risk unambiguously improves welfare when status quo relationships are assumed to remain unchanged, but may in fact increase the individual risk borne by the most risk-averse agents and decrease aggregate welfare, once the endogenous network response is taken into account. In Section 6, I describe an experiment to test the theory. Finally, I conclude. All technical details are relegated to the Appendices.

1.2 Background

The well-studied institution of sharecropping in Aurepalle, a village in Southern India which was among those sampled in the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT) dataset, provides a strong empirical context to fix the ideas of this model. Sharecropping is the practice of a farmer living on and cropping a plot belonging to a landowner, where the farmer pays rent as a *share* of realized profits (rather than as a fixed amount), and continues to be dominant in parts of the world today, e.g. rice farming in Madagascar (Bellemare (2009)) and Bangladesh (Akanda et al. (2008)).

Townsend and Mueller (1998) conduct a detailed survey of sharecropping relationships in Aurepalle. The modal type of arrangement in their sample is a landowner who hires a group of tenant farmers to work the land collectively, where the landlord actively participates in the farming process. The paper explicitly notes that there are many standard single-tenancy arrangements which were not sampled; however, the authors felt that the collective tenancy group was a reasonable approximation to the single-tenancy case. There is also strong evidence that risk concerns played a major role in the sharecropping relationship, as evidenced by implicit and explicit risk contingencies in the contracts.⁸

Notably, the paper finds that the *ex post* information sets of landlords and their tenants, along dimensions including realized harvest output, inputs, and reasons for crop failure, were found to coincide almost exactly. Communication and monitoring were found to be frequent and of high quality. Hence, although the abstraction away from moral hazard serves the purpose of focusing on the trade-off between use of *ex ante* and *ex post* insurance strategies, the model with this

⁸For more detail on and further examples of the influence of risk concerns on relationship formation, as well as the documented importance of risk attitudes in building risk-sharing relationships, please refer to the final Appendix.

assumption is still a close fit for empirically observed relationships.⁹

There is plenty of evidence that these landowners and farmers are risk-averse, and heterogeneous in their risk aversion. In a recent paper, Mazzocco and Saini (2012) develop a novel test and strongly reject a null hypothesis of homogeneous risk preferences in the ICRISAT dataset, for a very general class of preferences. Kurosaki (1990) finds evidence of substantial heterogeneity in risk aversion among ICRISAT landowners and farmers; furthermore, he does not find much evidence of heterogeneity in time preferences. Antle (1987) estimates the absolute Arrow-Pratt degrees of risk aversion for landowners and farmers in the Aurepalle data, and finds a mean risk aversion of 3.3 with a standard deviation of 2.6.

In addition to heterogeneity in risk attitude, individuals can choose to grow a variety of crops, where there is a great deal of heterogeneity in the riskiness of crop profitability. The riskiness of crop income derives from two sources, yield risk and price risk, where the yields and prices of different crops follow different probability distributions. Yield risk, for example, results in part from stochastic environmental factors, such as rainfall. Different crops react differently to idiosyncratic rainfall and soil conditions. Price risk results from farmers' inability to forecast what the world crop prices will be at the beginning of the growing season—this affects farmer's planting decisions.

In Aurepalle, farmers' crop portfolios largely consist of some combination of castor, pearl millet, cotton, pigeonpea, sorghum, chickpea, and paddy, and landowners and farmers who are heterogeneous in their risk aversion are observed to choose different portfolios of crops, portfolios of land plots, inputs, and farming methods (Townsend (1994), Lamb (2002)).

That the institution of sharecropping emerged in this environment has been the subject of much research. The literature has devoted substantial attention to understanding the feature of the share contract. In a well-known paper, Stiglitz¹⁰ in 1974 suggested that risk-averse landlords and tenant farmers adapt the incentive contracts of their employment relationships to accommodate their desire for risk protection, as a consequence of missing formal insurance. Townsend and Mueller (1998) explore the empirical relevance of a wide variety of ideas from the theory of mechanism design, such as monitoring and costly state verification. However, the question of which landlords match with which farmers, and why, is not well understood.

⁹Again, I model moral hazard explicitly in Wang (2012b). Moral hazard has interesting implications which are essentially orthogonal to the key economic insight of this paper, which focuses on how the heterogeneous tradeoff of risk management strategies across different risk partnerships influences the equilibrium matching pattern.

¹⁰Other papers which discussed similar ideas but differed in approach include Cheung (1968), Bardhan and Srinivasan (1971), and Rao (1971).

For further examples of informal insurance relationships, the importance of risk attitudes in building risk-sharing relationships, heterogeneity in risk attitudes, and heterogeneity in technology, please see the last Appendix.

1.3 The Model

In this section, I introduce a framework designed to analyze the equilibrium composition of partnerships between heterogeneously risk-averse people in a risky environment, where each partnership chooses which risky income stream they face as well as how they share that income.

The framework consists of the following elements:

The population of agents: the economy is populated by two groups of agents, $G1$ and $G2$, where $|G1| = |G2| = Z$, Z a finite, positive integer. All agents have CARA utility $u(x) = -e^{-rx}$, and are identical in every respect except for their Arrow-Pratt degree of absolute risk aversion r (and their group membership).¹¹ Assume that members r_1 of $G1$ are distinct in some unmodeled way from members r_2 of $G2$ (e.g. they differ in the type of capital they own, in time constraints, status, sex, etc.).

There are no distributional assumptions on the risk types in the economy.

The risky environment: a spectrum of risky projects $p > 0$ is available in the economy, where the returns R_p of project p are described by:

$$R_p = p + \varepsilon_p, \quad \varepsilon_p \sim N(0, V(p))$$

Hence, $R_p \sim N(p, V(p))$, where $V(p)$ describes the "variance cost" of a project with mean return p .

The function $V(p)$ has the following properties:

1. $V(0) = 0$ and $V(p) > 0$ for $p > 0$ (the unprofitable project is safe, and variance must be positive).

¹¹ Of course, in reality, types are multidimensional, and matching decisions are not exclusively based on risk attitudes. It is worth noting that the model can account for this. For example, individuals linked by kinship or friendship ties are more likely to know each other's risk types, and are more likely to trust each other, or be able to monitor and discipline each other. Hence, kinship and friendship ties would enter into this theory in the following way: an individual would first identify a pool of feasible risk-sharing partners, where this pool would be shaped by kinship and friendship ties, due to good information and commitment properties. Individuals would then choose their risk-sharing partners from these pools. This choice would be driven by risk attitudes, which is the focus of this model.

Thus, this theory can be thought of as addressing the stage of matching that occurs after pools of feasible partners have been identified.

2. $V'(0) = 0$, $V'(p) > 0$ for $p > 0$ (riskier projects have higher mean, but also higher variance).¹²
3. $V''(p) > 0$ ¹³.

Any project p requires the partnership of two agents, one from $G1$ and one from $G2$ ¹⁴. For example, a landlord must contribute land, and a farmer, expertise, in order for any crop to grow. Once matched, a pair (r_1, r_2) jointly selects a project. Assume that staying unmatched is disastrous for any agent, that is, an unmatched agent receives utility $-\infty$ ¹⁵. All projects are equally available to each possible pair, an agent can be involved in at most one project, and there are no "project externalities". That is, one pair's project choice does not affect availability or returns of any other pair's project.

There are no moral hazard considerations in this model. Please refer to Wang (2012b) for an explicit treatment of moral hazard in an endogenous matching problem.

Information and commitment: there is no information problem. All agents know each other's risk types and the risky environment (there is no disagreement about which risky projects are available, for example).

A given matched pair (r_1, r_2) observes the realized output of their partnership, and is able to commit *ex ante* to a return-contingent sharing rule $s(R_{p_{12}})$, where $R_{p_{12}}$ is the realized return of (r_1, r_2) 's joint project p_{12} . More precisely, $s(R_{p_{12}})$ specifies the income r_2 receives when the realized return is $R_{p_{12}}$, where $s : \mathbb{R} \rightarrow \mathbb{R}$ (there are no limited liability assumptions). In order to be *feasible*, the income r_1 receives must be less than or equal to $R_{12} - s(R_{p_{12}})$. Since all agents have monotonically increasing utility, r_1 's share will be equal to $R_{12} - s(R_{p_{12}})$.

Thus, an agent r_1 who is matched to r_2 under some sharing rule $s(R_{p_{12}})$ receives expected utility $E[-e^{-r_1(R_{p_{12}} - s(R_{p_{12}}))}|p_{12}]$, while the agent r_2 receives expected utility $E[-e^{-r_2 s(R_{p_{12}})}|p_{12}]$.

¹²The assumption that $V'(p) > 0$ is without loss of generality—we could begin by considering the entire space of project returns $N(p, V(p))$, where $V(p)$ is unconstrained. Then, any agent with concave utility would choose the project with lower variance, between two projects with the same mean. Tracing out the set of projects that agents would actually undertake leads us to $V'(p) > 0$.

¹³The assumption $V''(p) > 0$ is to ensure global concavity of an agent r 's expected utility in p and hence an interior solution for project choice.

¹⁴The fact that group size is bounded can be justified by, for example, costly state verification considerations. (For example, see Townsend and Mueller (1998) for a discussion of costly state verification and informal insurance in the Indian village of Aurepalle.) More generally, there is a sharp modeling tradeoff between allowing for a richer action space and allowing for richer group formation. Specifying that the group is of size two enables endogenous project choice and contract choice in each partnership.

¹⁵See Appendix 2 for a proof that the framework with joint project choice and infinite disutility from being unmatched is equivalent to the framework where each partner individually chooses a project and the returns are pooled, and individual rationality is a constraint that must be satisfied.

The equilibrium¹⁶: An equilibrium is:

1. A match function $\mu(r_1) = r_2$, where $\mu(\cdot)$ assigns each r_1 to at most one agent r_2 , and distinct people have distinct partners.

Moreover, the matching pattern described by $\mu(\cdot)$ must be stable. That is, it must satisfy two properties:

- (a) *No blocks*: no agent is able to propose an available project and a feasible sharing rule to an agent to whom she is not matched under μ , such that both agents are happier when matched with each other in this proposed arrangement than they are with the partners assigned by μ .
 - (b) *Individual rationality*: each agent must receive a higher expected utility from being in the match $\mu(\cdot)$ than from remaining unmatched.
2. A set of sharing rules and project choices, one sharing rule and project for each matched pair, such that no pair could choose a different sharing rule and/or a different project which leaves both partners weakly better off, and at least one partner strictly better off. In other words, the sharing rule and project chosen by a matched pair must be optimal for that pair.

Matching patterns: It will be helpful to introduce some matching terminology. Suppose the people in $G1$ and in $G2$ are ordered from least to most risk-averse: $\{r_1^j, r_2^j, \dots, r_Z^j\}$, $j \in \{1, 2\}$. Then "positive-assortative matching" (PAM) refers to the case where the i^{th} least risk-averse person in $G1$ is matched with the i^{th} least risk-averse person in $G2$: $\mu(r_i^1) = r_i^2$, $i \in \{1, \dots, Z\}$. On the other hand, "negative-assortative matching" (NAM) refers to the case where the i^{th} least risk-averse person in $G1$ is matched with the i^{th} most risk-averse person in $G2$: $\mu(r_i^1) = r_{Z-i+1}^2$, $i \in \{1, \dots, Z\}$. To say that the unique equilibrium matching pattern is PAM, for example, is to mean that the only μ which can be stable under optimal within-pair sharing rules and projects is the match function which assigns agents to each other positive-assortatively in risk attitudes.

A more detailed discussion of the elements of the framework follows the statement of results (with technical details relegated to the Appendix). I discuss the joint selection of a project by a partnership, the assumption of normally-distributed returns, the focus on partnerships rather than larger groups, and two-sided versus one-sided matching.

¹⁶ Existence is assured by Kaneko (1982).

1.4 Results

The heterogeneity of risk-aversion in agents makes this a model of matching under nontransferable utility. That is, the amount of utility an agent with preferences $u(x) = -e^{-r_1 x}$ (and hence risk type r_1) receives from consuming one unit of output x differs from the amount of utility an agent with preferences $u(x) = -e^{-r_2 x}$ (and hence risk type r_2) receives from consuming one unit of output x . Thus, we cannot directly apply the Shapley and Shubik (1962) result on sufficient conditions for assortative matching in transferable utility games.¹⁷

However, if a transferable utility representation for the model can be identified, then the Shapley and Shubik conditions for assortative matching can be applied. It will be helpful to review briefly their environment and result. Consider a population consisting of two groups of risk-neutral workers, where all workers have utility $u(c) = c$. Let a_1 denote the ability of workers in one group, and a_2 denote the ability of workers in the other group. The production function is given by $f(a_1, a_2)$, which can be thought of as: "the size of the pie generated by matched workers a_1 and a_2 ". Then, $\frac{d^2 f}{da_1 da_2} > 0$ is a sufficient condition for unique positive-assortative matching, and $\frac{d^2 f}{da_1 da_2} < 0$ is a sufficient condition for unique negative-assortative matching.

My approach here will be to identify the function in this model which is analogous to the Shapley and Shubik production function $f(a_1, a_2)$. In Proposition 1 below, I prove that *expected utility* is transferable in this model—instead of thinking about moving "*ex post*" units of output between agents, we should instead think about moving "*ex ante*" units of expected utility. I show that the sum of the certainty-equivalents $CE(r_1, r_2)$ of a given matched pair (r_1, r_2) is the analogy to the joint output production function in the transferable utility problem. The sum of the certainty-equivalents of a matched pair is "the size of the expected utility pie generated by matched agents r_1 and r_2 ", and sufficient conditions for positive-assortative and negative-assortative matching correspond to conditions for the supermodularity and submodularity of $CE(r_1, r_2)$ in r_1, r_2 .

More technically, expected utility is transferable in this model because the expected utility Pareto possibility frontier for a pair (r_1, r_2) is a line with slope -1 under some monotonic transformation.

Proposition 1 *Expected utility is transferable in this model.*

The proof is in Appendix 3.

¹⁷Legros and Newman (2007) provide a detailed and thoughtful discussion of the challenges and real world applications of solving for the equilibrium matching pattern when utility is not transferable.

Thus, conditions for the supermodularity and submodularity of $CE(r_1, r_2)$ in r_1, r_2 are sufficient conditions for unique PAM and NAM, respectively¹⁸. Finding these conditions requires characterizing the sum of certainty-equivalents generated by a matched pair. I sketch the important steps here (again, details can be found in Appendix 3).

To understand which individuals match in equilibrium and which project and sharing rule each of these equilibrium partnerships chooses, it is necessary first to understand the project and sharing rule chosen by a given matched pair (r_1, r_2) . Suppose r_1 and r_2 had already selected a project p . Then, conditional on r_1 ensuring r_2 some level of expected utility $-e^{-v}$, the optimal sharing rule solves:

$$\begin{aligned} \max_{s(R_p)} E \left[-e^{-r_1[R_p - s(R_p)]} | p \right] \quad s.t. \\ E \left[-e^{-r_2 s(R_p)} | p \right] \geq -e^{-v} \end{aligned}$$

Solving this program yields the following form for the optimal sharing rule:

$$s^*(R_p) = \frac{r_1}{r_1 + r_2} R_p - \frac{r_1}{r_1 + r_2} p + \frac{1}{2} \frac{r_1^2 r_2}{(r_1 + r_2)^2} V(p) + \frac{v}{r_2}$$

Note that the sharing rule is linear¹⁹, and the less risk-averse member of the partnership faces an income stream which is more highly dependent on realized project return.

Given the optimal sharing rule, and v , we can solve for the project a pair would choose. For any v , the pair agrees on project choice, since r_2 is ensured a fixed level of expected utility and is therefore indifferent over all projects. Hence, r_1 simply chooses the project that would make her the happiest, conditional on her having to ensure r_2 expected utility $-e^{-v}$. Furthermore, the pair agrees on the same project, *independent* of the "split of utility", v ²⁰.

Define a function $M(p)$, which describes the marginal variance cost of a small increase in mean

¹⁸Note that an alternative approach to characterizing sufficient conditions for assortative matching in models of nontransferable utility can be found in Legros and Newman (2007). Consider any four types $t_1, t'_1 \in G1, t_2, t'_2 \in G2$ in the economy, where $t_1 > t'_1$ and $t_2 > t'_2$. If, whenever agent t'_1 is indifferent between ensuring agent t_2 expected utility v and ensuring agent t'_2 expected utility u , agent t_1 *strictly prefers* (*strictly dislikes*) ensuring agent t_2 expected utility v to ensuring agent t'_2 expected utility u , then positive-assortative (negative-assortative) matching is the unique equilibrium.

These conditions are intuitive and an elegant generalization of the Shapley and Shubik result, but are often algebraically messy to use. The transferable expected utility approach proves to be much cleaner.

¹⁹Clearly, under moral hazard, incentive provision would also enter into the sharing rule—a more risk-averse individual is willing to pay a higher premium for insurance, but is also more expensive to incentivize. In Wang (2012a), I identify a framework of matching under one-sided moral hazard in which the equilibrium sharing rule is piecewise linear, and thus cleanly separates insurance and incentive provision.

²⁰Wilson (1968) showed that, as a consequence of CARA utility, any pair acts as a syndicate and "agrees" on project choice.

return from level p . That is:

$$M(p) = V'(p)$$

Then a partnership chooses a project by equating the marginal cost of increased variance with the marginal benefit of increased mean, weighted by their risk attitudes:

$$\begin{aligned} M(p^*) &= \frac{2(r_1 + r_2)}{r_1 r_2} \Leftrightarrow \\ p^* &= M^{-1} \left[\frac{2(r_1 + r_2)}{r_1 r_2} \right] \end{aligned}$$

Define the representative risk tolerance of a matched pair (r_1, r_2) :

$$\begin{aligned} \tilde{H}(r_1, r_2) &\equiv \frac{1}{r_1} + \frac{1}{r_2} \\ &= \frac{r_1 + r_2}{r_1 r_2} \end{aligned}$$

That is, we can think of a matched pair (r_1, r_2) acting as a single agent with CARA utility with risk tolerance $\tilde{H}(r_1, r_2)$. The expression for optimal project choice shows that a matched pair will choose

$$p^* = M^{-1} \left[2\tilde{H}(r_1, r_2) \right]$$

Since risk tolerance is the reciprocal of absolute risk aversion, we can also define the representative risk aversion of the matched pair:

$$\begin{aligned} \hat{H}(r_1, r_2) &\equiv \frac{1}{\tilde{H}(r_1, r_2)} \\ &= \frac{r_1 r_2}{r_1 + r_2} \end{aligned}$$

Now that we have characterized the project $p^*(r_1, r_2)$ and sharing rule $s(R_p | p^*(r_1, r_2), v)$ chosen by a matched pair (r_1, r_2) , we can characterize the partnership's sum of certainty-equivalents:

$$CE(r_1, r_2) = p^*(r_1, r_2) - \frac{r_1 r_2}{2(r_1 + r_2)} V(p^*(r_1, r_2))$$

That is:

$$CE\left(\tilde{H}(r_1, r_2)\right) = M^{-1}\left[2\tilde{H}(r_1, r_2)\right] - \frac{1}{2\tilde{H}(r_1, r_2)}V\left(M^{-1}\left[2\tilde{H}(r_1, r_2)\right]\right)$$

This expression provides both concrete and conceptual insight. Conceptually, we see that a matched pair can be thought of as a single CARA individual with risk tolerance \tilde{H} whose equilibrium welfare is captured by $CE(\tilde{H})$. This is because both individuals in a matched pair want to generate as large of an "expected utility pie" together as possible, regardless of the endogenously-determined split. Hence, the sum of certainty-equivalents for a given pair is a reasonable way to think about the welfare of that partnership, and of the individuals matched within it.

Concretely, we can see that more risk tolerant (higher \tilde{H}) matched pairs choose riskier (higher p) projects. In addition, the adverse impact of variance on $CE(\tilde{H})$ is smaller for higher \tilde{H} —that is, the size of the expected utility pie of a matched pair with representative risk tolerance \tilde{H} is less negatively-affected by the variance of a risky project for higher levels of risk tolerance.

It remains only to identify conditions on $CE(r_1, r_2)$ for its supermodularity and its submodularity in r_1, r_2 . Observe that:

$$\begin{aligned} \frac{d^2 CE(r_1, r_2)}{dr_1 dr_2} &= \frac{dCE}{d\tilde{H}} \frac{d^2 \tilde{H}}{dr_1 dr_2} + \frac{d^2 CE}{d\tilde{H}^2} \frac{d\tilde{H}}{dr_1} \frac{d\tilde{H}}{dr_2} \\ &= \frac{1}{r_1^2 r_2^2} \frac{d^2 CE}{d\tilde{H}^2} \end{aligned}$$

Hence, $CE(r_1, r_2)$ will be supermodular (submodular) in r_1, r_2 if and only if $CE(\tilde{H})$ is convex (concave) in risk tolerance \tilde{H} .

This leads to the main matching result on the fundamentals of the model.

Proposition 2 1. *A sufficient condition for PAM to be the unique equilibrium matching pattern is $M''(p) < 0$ for $p > 0$ (the marginal variance cost function is concave).*

2. *A sufficient condition for NAM to be the unique equilibrium matching pattern is $M''(p) > 0$ for $p > 0$ (the marginal variance cost function is convex).*

3. *A sufficient condition for any matching pattern to be sustainable as an equilibrium is $M''(p) = 0$ for $p > 0$ (the marginal variance cost function is linear).*

(Technical details of the proof can be found in Appendix 4.)

Note that this result shows that the equilibrium matching pattern is independent of the distributions of risk types in the economy, which makes its application very flexible.²¹

What is the intuition behind the matching result? We know from the literature the intuition behind negative-assortative matching when agents lack income-smoothing tools—in that case, it is the least risk-averse guy in $G1$ who is "willing to place the highest bid" for the most risk-averse agent in $G2$. This is because the least risk-averse agent is the most willing to provide insurance, while at the same time the most risk-averse agent has the highest willingness to pay for good insurance. This "bidding" turns out to have a monotonic property, in that once the least risk-averse agent in $G1$ and the most risk-averse agent in $G2$ are matched and are "removed from the pool", the least risk-averse of the remaining agents in $G1$ is then matched with the most risk-averse of the remaining agents in $G2$, and so on.

My result shows that accounting for the interaction of income-smoothing with consumption-smoothing leads to the emergence of *both* extremal matching patterns. Loosely, a risk-averse individual prefers a partner unlike herself for consumption-smoothing purposes (a "gains from trade" matching motivation), but prefers a partner like herself for income-smoothing purposes (a "similarity of perspective" matching motivation). In a model with both consumption- and income-smoothing, playing the role of informal insurer isn't as straightforward for the less risk-averse agents—the consumption-smoothing that a less risk-averse agent offers to a more risk-averse partner comes at a cost. Namely, the insurance that a less risk-averse agent provides within the relationship affects the choice of income stream—if the less risk-averse agent did not have to bear most of the consumption risk, she would have preferred to choose a riskier income stream with higher mean return.

When the marginal variance cost is increasing in expected return, and increasing more rapidly for higher levels of expected returns, the benefit of playing the role of informal insurer dominates the costs for the less risk-averse agents, because optimal consumption-smoothing does not come at much of a cost—it does not really involve a sacrifice in the choice of income stream, since the less risk-averse agents also prefer to stick to safer projects. In this environment, there is little friction between consumption-smoothing and income-smoothing for all agents across all partner types, and thus the "gains from trade" consumption-smoothing motivation dominates, resulting in

²¹These sufficient conditions could be weakened slightly by introducing slight dependence on the distributions of risk types, only through the supports: the concavity, convexity, and linearity of $M(\cdot)$ need only to hold on the domain $\left[\frac{2}{\max(r_1)} + \frac{2}{\max(r_2)}, \frac{2}{\min(r_1)} + \frac{2}{\min(r_2)}\right]$, and not on \mathbb{R}^+ .

unique negative-assortative matching.

However, when the marginal variance cost is increasing in expected return, but increasing more *slowly* for higher levels of expected returns, the role of informal insurer becomes less appealing for a less risk-averse agent. Now, when the less risk-averse agent shoulders most of the risk in a relationship with a more risk-averse agent, she must make a sacrifice in the choice of income stream—were she to be slightly more insured herself, she would start a project with a much higher expected return, since marginal variance cost is concave in expected returns. This highlights a natural and important constraint of informal insurance which is typically not emphasized: each risk-averse individual can share risk only with *other risk-averse individuals*. Hence, in this environment, the less risk-averse agents experience friction between consumption-smoothing and income-smoothing when partnered with more risk-averse agents, and this friction diminishes when they pair up with fellow less risk-averse agents instead. This causes the less risk-averse agents to choose entrepreneurial pursuits with other less risk-averse agents, over informal insurance relationships with more risk-averse agents, and unique positive-assortative matching results.

Finally, when the marginal variance cost is increasing linearly in expected return, the total sum of certainty-equivalents is the same across all possible matching patterns. Partnerships of different risk compositions do generate different pair-specific certainty-equivalents, but all risk types in $G1$ have the same *strength of preference* between any two individuals in $G2$ (and vice versa). Hence, any matching pattern is sustainable as an equilibrium.

It is worth asking what happens in the model when the income-smoothing channel is shut down, and what happens when the consumption-smoothing channel is shut down. Shutting down the income-smoothing channel effectively corresponds with the Chiappori and Reny (2006) and the Schulhofer-Wohl (2006) settings, so it is reassuring that in this setting, in the absence of project choice, negative-assortative matching arises as the unique equilibrium. Furthermore, when instead the consumption-smoothing channel is shut down (by holding the division of output fixed at a constant across all possible pairs), the unique equilibrium matching pattern is positive-assortative. (See Appendices 5 and 6 for details.)

Can this theory be falsified empirically? An obvious challenge of directly checking the concavity or the convexity of the marginal variance cost function according to the condition identified in Proposition 2 is that we may not observe the variance cost function in the data. For instance, one obstacle is that even a bad estimator of the variance of a project requires more than one observation of its returns, and good dynamic data, especially in developing countries, is difficult to

gather. However, other elements of the model are more conducive to data collection. For example, existing papers have collected data on the risk types of individuals, the linkages between these individuals, and some proxy of mean income, e.g. a household's realized income (in the absence of any other information about the distribution of returns, and for a large sample of households, this will be a reasonable proxy for the mean).

The following two results will be helpful for more concretely linking the intuition of the matching equilibrium to the data.

Proposition 3 *The mean return of optimal project choice $p^*(r_1, r_2)$ if (r_1, r_2) are matched is supermodular ($\frac{d^2 p}{dr_1 dr_2} > 0$) in r_1, r_2 iff $M''(p) < 0$, and is submodular ($\frac{d^2 p}{dr_1 dr_2} < 0$) in r_1, r_2 iff $M''(p) > 0$.*

Proposition 4 *The mean return of optimal project choice $p^*(r_1, r_2)$ for a pair (r_1, r_2) is convex in representative risk tolerance $\tilde{H}(r_1, r_2)$ iff $M''(p) < 0$, and is concave in $\tilde{H}(r_1, r_2)$ iff $M''(p) > 0$.*

(Recall that $\tilde{H}(r_1, r_2) \equiv \frac{1}{r_1} + \frac{1}{r_2}$.)

(See Appendix 4 for the proofs.)

In words, the first corollary tells us that the sufficient condition for unique positive-assortative matching corresponds exactly to the supermodularity of the mean return of a matched pair's optimal choice of project in the risk types of the matched pair, while the sufficient condition for unique negative-assortative matching corresponds exactly to the submodularity of the mean return of a matched pair's optimal choice of project in the risk types of the pair.

However, we cannot directly test for this relationship, because we only observe *equilibrium* projects.

To be more clear about this, consider the following example. Suppose $G1 = \{r_1^A, r_2^A, r_3^A\}$ and $G2 = \{r_1^B, r_2^B, r_3^B\}$, where $r_1^{A,B} < r_2^{A,B} < r_3^{A,B}$.

Then, suppose the underlying marginal variance cost function is such that $M''(p) < 0$. Proposition 1 says that this means we observe agents positively-assorting along risk attitudes. But we do not observe the marginal variance cost function $M(p)$, or the variance function $V(p)$. We only observe mean project returns in equilibrium, $p(r_1^A, r_1^B)$, $p(r_2^A, r_2^B)$, and $p(r_3^A, r_3^B)$, as well as the risk types of agents in $G1$ and $G2$, and the positive-assortative linkages between $G1$ and $G2$.

If we could show supermodularity of the mean of optimal joint project choice of a matched pair in r_i^A, r_j^B , then this would be evidence supportive of the theory. But we do not observe $p(r_i^A, r_j^B)$ for $i \neq j$ —that is, we do *not* observe the projects that individuals who are unmatched in equilibrium

would have chosen had they been forced to match. So, we cannot use this result to see test whether or not the data supports the theory.

Fortunately, Proposition 3 implies Proposition 4, which is testable with data we are likely to have. We can use our observed matchings (r_1^A, r_1^B) , (r_2^A, r_2^B) , and (r_3^A, r_3^B) to calculate the representative risk tolerance $H_i \equiv \tilde{H}(r_i^A, r_i^B)$ for each matched pair. Then, we can regress $p_i \equiv p(r_i^A, r_i^B)$ on a constant, as well as \tilde{H}_i and \tilde{H}_i^2 (to use a crude polynomial approximation). If the regression shows that p_i is convex in \tilde{H}_i , that is, the coefficient on \tilde{H}_i^2 is positive, then this is evidence supportive of the theory, since we have established the following: (a) individuals are matched positive-assortatively in risk attitudes in the data, and (b) the mean returns of equilibrium projects are convex in the representative risk tolerance of the matched pairs.²² We know from the theory that (b) is equivalent to the original matching condition of concavity of the marginal variance cost function $M''(p) < 0$.

Hence, it is clear that there are concrete, observable differences between environments in which PAM arises as the unique matching equilibrium, and environments in which NAM arises as the unique matching equilibrium. The falsifiability condition is constructed by exploiting these differences. Loosely, equilibrium project choice varies less and representative risk types are more clustered in environments where NAM arises than in environments where PAM arises.

Finally, can we say anything about efficiency?

Proposition 5 *The equilibrium maximizes the sum of certainty-equivalents, and is Pareto efficient.*

Proposition 6 *The equilibrium maximizes total mean project returns (conditional on a matched pair choosing its project optimally).*

Both of these propositions follow straightforwardly from the conditions for positive- and negative-assortative matching.

Since the sum of certainty-equivalents is a social welfare function, and the equilibrium maximizes this sum, it must be Pareto efficient. Intuitively, if an agent could be made better off without making his partner worse off via a change in sharing rule or project choice, then that sharing rule or project choice was not optimal in the first place and thus not an equilibrium. If an agent could be made

²²In some sense, this condition is necessary but not entirely sufficient. What we really want is to confirm convexity or concavity of the marginal variance cost function over the entire interval between the lowest possible partnership risk tolerance and the highest possible partnership risk tolerance. Instead, we observe only points within this interval, corresponding to the risk tolerances of matched pairs. However, for a large enough population, this is a reasonable approximation.

better off without making anyone worse off by switching partners, then that switch would have occurred and the original assignment would have violated the first stability criterion.

However, while the sum of certainty-equivalents of individuals within a pair is a reasonable way of thinking about the welfare of those individuals and that partnership, the unweighted sum of certainty-equivalents across all pairs is not the right way of thinking about aggregate welfare in this economy. To understand various reasons for this, first recall the expression for the certainty-equivalent of a matched pair with representative risk tolerance \tilde{H} :

$$CE(\tilde{H}) = p - \frac{1}{2\tilde{H}}V(p)$$

Note that at every possible project p , the certainty-equivalent of a pair with higher risk tolerance \tilde{H} will be larger than a less risk-tolerant pair. This is precisely because a less risk-tolerant individual dislikes risk more—she requires a smaller amount with certainty to receive the same expected utility she does from facing a risky project p . This implies that if the population consisted of two pairs, one more risk tolerant than the other, and a social planner cared about the sum of certainty-equivalents across pairs but could only allow one pair to start a project (leaving the other pair to receive $-U$ in the absence of a project), then the social planner would always choose to give the project to the more risk tolerant pair. But this is unsatisfying, as there is no inherent social value in doing this. That is, there is no reason it should be viewed as worse from society's perspective to give a project to an individual who requires a smaller amount to be indifferent between receiving that certain amount and facing the project, versus an individual who requires a larger amount. Moreover, in this model, transfers across pairs are not allowed, so an unweighted sum across pairs relates pairs to each other in an artificial way.

Hence, a superior approach to policy evaluation in this setting is to ask whether a policy generates a Pareto improvement. Alternatively, the use of a Rawlsian social welfare function would also be appropriate, as the welfare of the most risk-averse agents is a natural priority in this setting.

The second proposition tells us that, in a given environment, any movement of the matching pattern away from the equilibrium results in a decrease in total expected project returns, as long as a matched pair chooses its project optimally.

1.4.1 Support in the Literature

Is there support for this theory in existing empirical work? The data from the experiment of Attanasio et al. (2012) and the sharecropping data from Akerberg and Botticini (2002) suggest that the answer is yes. (In Section 5, I describe the ideal experiment to test this theory.)

Attanasio et al. (2012) run a unique experiment with 70 Colombian communities. They elicit risk attitudes by privately offering subjects a choice of gambles in the first round, which vary in riskiness (higher mean projects come at the cost of higher variance). In addition, they gather data on pre-existing kinship and friendship networks. Kinship and friendship ties matter for two important reasons: first, individuals are likely to know the risk attitudes of family and friends, and unlikely to know the risk attitudes of strangers. Second, individuals are likely to trust family and friends over strangers (renegades cannot be detected in this experiment).

Individuals are able to choose what gamble they face, but the sharing rule for each risk-sharing group is fixed at equal division. In other words, individuals are able to manage risk *ex ante*, but not *ex post*.

Attanasio et al. study a simplified model of their experimental setting to identify theoretical predictions: individuals have CARA utility and have either *low* or *high* risk aversion (type is binary). In addition, they match pairwise and can choose which risky project to undertake, but the sharing rule is fixed at equal division. Finally, they can be either familiar or unfamiliar with each other, where familiar individuals know each other's risk types and can perfectly commit to each other. Attanasio et al. show that familiar individuals should match positive-assortatively in risk attitude: people with high risk aversion partner with each other, and people with low risk aversion partner with each other. This is verified by the experiment, where family and friends are observed to group together with greater likelihood, and conditional on this, to assort positively in risk attitudes

The model in this paper enables us to understand these experimental results in a more general context. For example, the half-half split of project returns in the model is key to the proof of the positive-assortative matching result in Attanasio et al. (2012), not because of the equality, but because no constant transfer is permitted. This means that utility is completely nontransferable, and outside options are not endogenously determined. The model would be very different if the form of the sharing rule were $s(R_p) = \frac{1}{2}R_p + b$, where individuals can choose b . My results show that it is precisely the trade-off between income-smoothing and consumption-smoothing across pairs which

is essential for the equilibrium match. In particular, positive-assortative matching arises uniquely when the consumption-smoothing channel is shut down (see Appendix 6 for details).

I seek support in the data from the Attanasio et al. experiment in the following manner. I focus on the subset of individuals reporting at least one family/friendship tie with their matched group members, since these "familiar" individuals know each other's risk attitudes and are able to commit to a return-contingent sharing rule, and these are key features of my model. Note that family and friendship ties can be incorporated into the framework of this paper by thinking about such ties as entering the matching process in a "period 0" stage, where individuals first identify pools of feasible risk-sharing partners. Family and friends tend to fall into this pool for informational and commitment reasons. Then, the choice of partner from this pool is driven by risk attitudes, and is determined by the forces described in this paper.

The choice of gambles is described below²³. Each gamble had a low payoff and a high payoff (units are Colombian pesos), where the probability of each payoff was $\frac{1}{2}$ across all gambles²⁴. Riskier gambles had lower low payoffs but higher high payoffs. The paper provides the ranges of risk aversion corresponding to each gamble choice; I list the lower bound of each range²⁵:

Gamble	Mean	Variance	Risk Aversion
1	3000	0	$[7.49, \infty)$
2	4200	2,250,000	$[1.73, 7.49)$
3	4800	5,760,000	$[0.81, 1.73)$
4	5400	12,960,000	$[0.46, 0.81)$
5	6000	25,000,000	$[0, 0.46)$

Thus, it is possible to characterize (approximately) the variance function $V(p)$, which describes

²³The standard deviations reported in the table in Attanasio et. al. (2012) are the experimental standard deviations, and not the analytical standard deviations based off the equally likely low and high payoffs of each gamble. The discrepancy results from altruism on the part of the experimenters, who were found to have a tendency to give the subject the high payoff even when the subject lost the gamble. Here, I consider the analytical variance, since Attanasio et. al. state that subjects were unaware of the bias in probabilities during their private, one-on-one risk elicitation rounds. Hence, subjects should have thought they were facing the analytical variance.

²⁴The distribution of these gamble returns is clearly not normal. This analysis is merely meant to be supportive of the theoretical results.

²⁵I chose the lower bound primarily because the risk type interval corresponding with the safest gamble is $[7.49, \infty)$, and choosing any other point in that large interval seems particularly arbitrary.

Additionally, the lower bound of the interval corresponding to the fifth gamble is 0; I approximate this with 0.0005.

I drop the sixth and riskiest gamble, because the fifth and sixth gambles have the same mean, but the sixth gamble has higher variance. Thus, only risk-loving individuals should choose the sixth gamble, and in my model, I consider only risk-averse people. About 15% of people chose the risk-loving gamble in the risk type elicitation round.

the variance of a project with mean p . Convexity or concavity of the marginal variance cost function does not determine equilibrium matching here, since the shutdown of the consumption-smoothing channel ensures positive-assortative matching, but it will influence the mean returns of projects chosen by matched pairs in equilibrium.

Hence, I take the following steps. The observation of positive-assortative matching among familiar individuals is the first point of support, since this is also what is predicted by the more general theory. Second, I seek confirmation that members of matched groups chose similar projects. Third, I solve for the closest power function approximation of $V(p)$. Finally, I test the data to see if the behavior of the mean returns of equilibrium projects in risk tolerance aligns with the behavior predicted by the theory.

The second point is easily verified: members of matched groups did indeed choose similar projects. The mean difference in second-round gamble choice within groups of familiar individuals is 0, and the modal difference is also 0.

The closest power function approximation of the variance function is $V(p) = (p - 3000)^{2.1}$. We know from the theory that this implies the mean returns of the projects chosen in equilibrium should be slightly concave in risk tolerance. More specifically, if it were the case that $V(p) = (p - 3000)^2$, the mean returns would be predicted to be linear in risk tolerance. If the variance function is more convex than a quadratic function, the mean returns should be concave in risk tolerance, while if the variance function is less convex, the mean returns should be convex. Since the power of the function is 2.1 and not 2, the observed mean returns should be "slightly" concave in risk tolerance.

I regress the mean returns of the projects chosen on risk tolerance and squared risk tolerance. A positive coefficient β_3 would indicate convexity, while a negative β_3 would indicate concavity²⁶:

$$p_i = \beta_1 + \beta_2 \left(\frac{1}{r_i} \right) + \beta_3 \left(\frac{1}{r_i} \right)^2 + \varepsilon_i$$

The OLS estimates are reported in Table 1:

²⁶There are more precise ways to examine data for convexity or concavity (e.g. local linear regression or nonparametric methods). The purpose here is not to develop the optimal econometric technique for convexity-testing, but rather to describe an approach to the problem and demonstrate its implementability by executing the approach in a simple, intuitive way. Here, this is done by approximating $V(p)$ with a polynomial.

Table 1: Testing for Concavity of Mean Project Returns in Risk Tolerance

Regressor	OLS Estimates
Sq. risk tolerance	-0.001*** (-60.73)
Risk tolerance	2.39*** (2.75)
Constant	5169.18*** (180.91)

Notes: This table reports the OLS estimates from a regression of mean project returns on squared risk tolerance, risk tolerance, and a constant. Data from Attanasio et al. (2012).

*significant at 10%; **significant at 5%; ***significant at 1%

In line with the theoretical prediction, we see that mean project returns are concave in the rate of risk aversion, and indeed they are even *slightly* concave (close to linear). Moreover, individuals with a higher risk tolerance (less risk-averse individuals) choose riskier (higher mean) gambles.

Hence, the theory finds preliminary support in these experimental results: positive-assortative matching in risk attitudes is observed among individuals who know each other's risk attitudes and are therefore able to trust and commit, and who are only able to choose what risk to share, not how to share a given risk. Furthermore, individuals who choose to match select similarly risky gambles, and the mean returns of the chosen gambles are slightly concave in risk tolerance. This is as predicted by the theory, since the variance function is approximately $V(p) = (p - 3000)^{2.1}$, just a bit more convex than the quadratic threshold.

A second paper which offers empirical support for this theory is Akerberg and Botticini (2002). They provide evidence of endogenous matching between landowners and sharecroppers in medieval Tuscany, motivated by risk concerns. They assume that landowners exogenously owned crops of varying riskiness (the safe crop of cereal, the risky crop of vines, and mixtures of the two), and studied the matching of tenants with differing risk attitudes to different landowners/crop plots. That is, they assume that individuals were able to manage risk *ex post*, but were restricted in their ability to manage risk *ex ante*. In their data, they observe that fixed-rent (residual claimancy) contracts were associated with the safer crop of cereal, while share contracts were associated with the riskier crop of vines. In addition, they observe that less risk-averse agents were matched to cereals, while

more risk-averse agents were matched to vines. They then provide empirical support for their endogenous matching conjecture that less risk-averse agents worked under fixed-rent contracts for more risk-averse landlords who owned the safer crop of cereal, while more risk-averse agents worked under share contracts for less risk-averse landlords who owned the riskier crop of vines.

But the data also shows that the types of crops cultivated differed starkly across two regions of Tuscany they studied: San Gimignano and Pescia. The following table shows the number of pure vine plots, pure cereal plots, and mixed plots cultivated in Pescia and in San Gimignano:

	<i>Pescia</i>	<i>San G.</i>
	# plots	# plots
Vines (risky)	47	6
Mixed	4	127
Cereals (safe)	178	17

The difference in the equilibrium crop mix is striking. In particular, "mixed" is not a third crop in and of itself, but a mixture of the risky crop of vines and the safer crop of cereals (likely differing mixtures, but the degree of mixing is not recorded). Hence, it is puzzling why, even though vines and cereals were available to be grown in both areas, the land plots of San Gimignano were overwhelmingly mixtures, while mixtures were almost entirely absent in Pescia. How does this setting fit into the framework of this paper?

The theory suggests that the agroclimactic differences between Pescia and San Gimignano may be important. These two regions are located fairly close together (about 100 km apart), and are both Tuscan regions, so they shared the same governance and were unlikely to differ in norms. However, Pescia experiences higher mean rainfall than San Gimignano (Dalla Marta et al. (2010)). A primary reason higher mean crops have higher variance of yield is that they are more sensitive to rainfall—given a high and steady amount of rainfall, the crops can yield a substantial harvest, but any shortage will lead to blight. Safer crops can produce a positive yield even with low levels of rainfall, but added rainfall is unlikely to increase this yield significantly.

Hence, crops which were less sensitive to rain (such as cereals and legumes) may have had similar distributions of yields in both Pescia and San Gimignano, while crops which were more sensitive to rain (such as vines and olives) were likely to have had higher variance of yield in San Gimignano than in Pescia. In the language of this model, the marginal variance cost function of crops with mean p in San Gimignano was likely to have been convex in p ($M''(p) > 0$), while the

marginal variance cost function of crops with mean p in Pescia was likely to have been concave in p ($M''(p) < 0$).

Thus, while the equilibrium choice of crop portfolio in both regions seems to have encompassed the spectrum between cereal crops and vines and excluded crops such as legumes and olives, the theory suggests that landowners and farmers in Pescia matched positive-assortatively and chose the endpoints of this spectrum (pure cereals and pure vines), while landowners and farmers in San Gimignano matched negative-assortatively and chose the middle points of this spectrum (mixtures of crops and vines). These are in fact the cropping patterns we observe in the data.

1.4.2 Discussion of the Model

The specification that agents jointly start a project and hence never wish to remain unmatched may seem restrictive, but in fact nests the following framework. Suppose each agent chooses her own risky income stream. Subsequently, risk-sharing partnerships are formed, and a matched pair pools their realized incomes and splits the pooled amount according to some agreed-upon sharing rule. An equilibrium matching must satisfy both the stability requirement, and an individual-rationality (*IR*) requirement: the expected utility any agent receives in her partnership must exceed her utility from being alone.

I show rigorously that my model nests this framework in Appendix 2. The idea behind the equivalence is the following: in this framework where individuals choose their own risky income streams, the autarkic sharing rule is always an option for a matched pair. Hence, *IR* is always satisfied—two individuals can always match and agree to a sharing rule of the form, "Each person keeps her own realized income." There are no explicit costs to matching, so no one remains unmatched.

Furthermore, since a pair of agents pools their realized incomes, couples are essentially choosing their preferred pooled income stream when they match in equilibrium (if income streams are independent, this is simply the convolution of individual income distributions). But this is equivalent to couples explicitly jointly choosing a risky project from some spectrum, and assuming that no agent stays unmatched. This latter specification is much more natural for an analysis of income-smoothing, consumption-smoothing, and endogenous matching.

In this model, I also assume that returns are normally-distributed. The normal distribution has a "representational convenience", in that it has the feature that the only nonzero cumulants²⁷

²⁷Recall that the cumulant-generating function is the log of the moment-generating function. The first and second

are the mean and the variance. A natural concern is that this "representational convenience" is driving the result.

I solve a generalization of the model under the assumption that returns follow an arbitrary symmetric distribution, and show that normality is not driving the result. The normal distribution does generate unusual tractability: the mean-variance characterization yields a natural parameterization of the risky project space, so that the function $V(p)$ captures entirely the "cost" of gains in expected return in terms of variance, since there are no higher order cumulants. But consideration of a generalized model shows that the key underlying determinant of equilibrium matching is still a cost-benefit trade-off across the spectrum of risky projects, where this trade-off is simply less concise than it was before: the "cost" of gains in expected return is now captured by the sum of all the higher order cumulants, not just the variance. But the sum of all the higher-order cumulants essentially acts as a "generalized variance" in this context—the sufficient condition for matching is still a parametric condition which can be calculated. Thus, there is no loss in economic insight by specifically assuming normally-distributed returns. Please refer to Appendix 1 for more details and results for the generalized model.

A third piece of the model is the restriction to groups of size two. Costly state verification provides an empirical justification for a bound on group size. However, the assumption of partnerships is important for the theory. It is indeed a restriction, but there is a very sharp modeling trade-off between richness of action space and richness of network formation. Focusing on a network of pairs allows me to optimize over the whole space of sharing rules and projects, so I can analyze how the trade-off between income-smoothing and consumption-smoothing within a pair influences the composition of the equilibrium network. This would not be possible if I allowed for arbitrary kinds of match formation.

A final piece of the model is the two-sidedness of the matching problem. Why not one-sided matching? The primary reason for this is technical: under one-sided matching, "negative-assortative" and "positive-assortative" no longer uniquely identify matching patterns. That is, when there are two distinct groups of agents, e.g. $G1 = \{1, 2, 3\}$ and $G2 = \{3, 4, 5\}$, the positive-assortative matching is clearly $(1, 3), (2, 4), (3, 5)$. However, if there were only one group of agents, $G = \{1, 2, 3, 3, 4, 5\}$, then $(1, 3), (2, 3), (4, 5)$ is also a positive-assortative matching pattern. Moreover, theoretically there may be problems with existence in one-sided matching problems. These reasons are economically superficial—it is clear that the assumption of two-sided matching is merely

cumulants of any distribution are the mean and the variance.

for convenience, and does not affect the insights of the theory.

1.5 Policy

In this section, I show how to use the theory to analyze the aggregate and distributional welfare impacts of a hypothetical policy. The natural social welfare function in this setting is the total sum of certainty-equivalents (see the discussion at the end of the Results section).

Using the theory concretely requires consideration of a few practical issues. For example, the potential for the two groups in a population to be of different sizes (e.g. there may be more tenant farmers than landowners) is addressed in the following subsection.

1.5.1 Differently-sized groups

The assumption of equally-sized groups is unlikely to hold in real-world settings. In the benchmark model, it is a convenient and relatively harmless assumption to make, because the goal was to understand the theoretical impact of the scope of income-smoothing and consumption-smoothing informal insurance strategies within a partnership on the risk composition of equilibrium partnerships, and this insight does not depend on group size.

However, if we want to use this framework to evaluate policies, we need to account for the possibility of differently-sized groups, as this could conceivably affect, among other things, the magnitude of the policy impact. To consider an extreme example, if one group consists only of a single person, while the other group consists of some large number of people, then the individual in the first group may change partners in response to the policy, but most people will be unaffected. Therefore, to realistically think about policy, we need to understand what would happen in this model if the two groups were of different sizes.

Lemma 7 *If $|G1| < |G2|$ ($|G1| > |G2|$) the most risk-averse excess agents of $G2$ ($G1$) will be unmatched. The remaining agents are matched according to the main result (Proposition 2).*

A proof is provided in Appendix 7.

Hence, we see that if there are more tenant farmers than landlords, for example, the most risk-averse tenant farmers will be unmatched, and we can use the conditions from Proposition 2 to think about the endogenous matching between the remaining farmers and landlords.

1.5.2 Crop price stabilization and formal insurance

Price stabilization, particularly of crops, is frequently proposed by governments as a tool for alleviating the substantial risk burden shouldered by the poor (Knudsen and Nash (1990), Minot (2010)). Farmers face a large amount of yield and price risk. Crop price stabilization should therefore reduce the income risk of farmers, and should particularly benefit the most risk-averse, poorest agents (Dawe (2001)).

A key contribution of my framework is to illuminate the importance of accounting for the endogenous informal risk-sharing network response when analyzing the costs and benefits of such a policy, a point that, to my knowledge, has not yet been made precise in the literature.

How should crop price stabilization be modeled here? Some crops have higher price risk than others. In particular, one channel leading to differences in price risk is differences in yield risk. When yield is more uncertain (because the crop is more sensitive to weather shocks and growing conditions), the world price is more uncertain and expectations about what that price will be at the start of the growing season are correspondingly noisier. On the other hand, when a crop is very robust to weather shocks and growing conditions, so that yield fluctuates very little, the world price also fluctuates less, and expectations about that price are much sharper at the start of the growing season. Hence, in this exercise, we can think of risky crops as having high mean and variance of yield, and noisy expectations about price, while safe crops have low mean and variance of yield, and sharp expectations about price.

Now, consider a setting where higher mean crops come at an incredibly high cost of price and yield risk. In particular, the marginal variance cost is convex in mean returns. Suppose the government wishes to encourage producers to grow higher mean crops, and implements a variable tariff that sets a price band around the prices of the riskiest (highest mean) crops. This policy causes aggregate risk to fall, since $V(p)$ is weakly smaller for all $p > 0$, and decreases the variance $V(p)$ by a larger amount for higher p crops.

It will be helpful to employ a simple functional form characterization to analyze this effect. Specifically, suppose that in the status quo, the profits of a crop with mean p followed this distribution: $\pi_p \sim N(p, p^{N_1})$. Following the stabilization policy, the distribution of profits is $\pi_p \sim N(p, p^{N_2})$, where $N_1 > N_2$: that is, $V(p)$ fell for each p , and became less convex.

We know from Proposition 2 that if $R_p \sim N(p, p^N)$, the unique equilibrium matching pattern is NAM if $N > 2$ and the unique equilibrium matching pattern is PAM if $N \in (1, 2)$. Hence, suppose

$N_1 > 2$, and consider a price stabilization policy which reduces the convexity of variance so that $N_2 \in (1, 2)$.

Note that using the power function as a functional form for $V(p)$ has one small drawback. We want to analyze a policy that reduces the variance of every project (reduces the riskiness of an environment), and particularly reduces the variance of the riskiest projects, which makes "decreasing N " a natural choice for representing this policy. However, when N falls, the variance of the projects $p \in (0, 1)$ actually increases. To focus on our policy analysis, it is most straightforward to simply assume that the population of risk types in $G1$ and $G2$ is such that no possible pair ever wishes to undertake a project $p \in (0, 1)$. So, assume:

$$\frac{\bar{r}_1 + \bar{r}_2}{\bar{r}_1 \bar{r}_2} \geq \frac{N_1}{2}$$

where \bar{r}_1 is the most risk-averse agent in $G1$, and \bar{r}_2 is the most risk-averse agent in $G2$. This is simply for convenience, and places no substantive restrictions on the intuition or the policy analysis, especially as my matching results are distribution-free. Now, decreasing N unambiguously reduces the riskiness of the project environment and captures our policy experiment.

There are two effects of altering the shape of $V(p)$: first, there is a pure effect from decreasing the riskiness of an environment (which causes any matched pair to select a higher p project), and second, there is the endogenous network response resulting from it, which then further affects welfare.

We know the expression for a partnership's sum of certainty-equivalents:

$$\begin{aligned} CE(r_1, r_2) &= p^*(r_1, r_2) - \frac{1}{2} \frac{r_1 r_2}{r_1 + r_2} V(p^*(r_1, r_2)) \\ &= V^{N-1} \left(\frac{2(r_1 + r_2)}{r_1 r_2} \right) - \frac{1}{2} \frac{r_1 r_2}{r_1 + r_2} V \left(V^{N-1} \left(\frac{2(r_1 + r_2)}{r_1 r_2} \right) \right) \end{aligned}$$

In this example, we consider $V(p) = p^N$, so:

$$CE(r_1, r_2) = \left[1 - \frac{1}{N} \right] \left[\frac{1}{N} \frac{2(r_1 + r_2)}{r_1 r_2} \right]^{\frac{1}{N-1}}$$

We know from Proposition 2 that if $N_1 > 2$ falls to $N_2 \in (1, 2)$, the unique equilibrium matching pattern switches from negative-assortative to positive-assortative.

First suppose that a policy decreases N , so that the variance of every project $p \geq 1$ is reduced, and the variance of higher mean projects is particularly reduced, but that partnerships remain

unchanged.

Lemma 8 *A policy which reduces the variance of every available project is a strict Pareto improvement if the composition of partnerships is unaffected.*

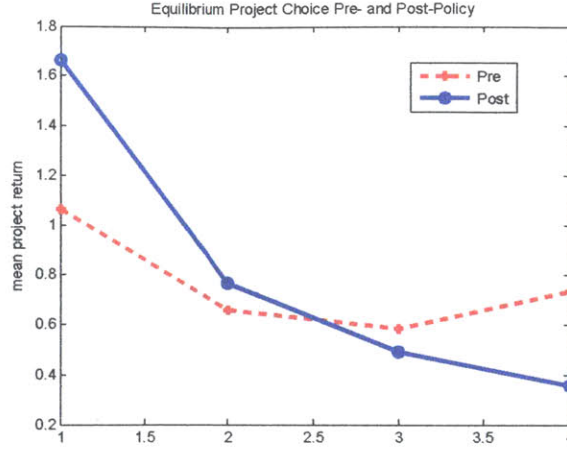
The proof is intuitive and very general: suppose that the Z matched partnerships in the status quo undertook projects p_1, \dots, p_Z . Following the introduction of a policy which decreases the variance of every project, each partnership has the opportunity to stay on its original project, or to switch projects. If a partnership retains its original project, it is strictly better off, since the project has the same mean, but a lower variance, and all individuals are risk-averse and thus dislike variance. If a partnership switches to a different project, then by revealed preference, it must be even better off facing the new project than facing the old project with decreased variance. But this means that each partnership is strictly better off.

Now, suppose we account for the potential endogenous re-formation of partnerships triggered by the policy. A numeric example clearly illustrates the mechanism by which this policy could negatively affect welfare, purely through the channel of endogeneity.

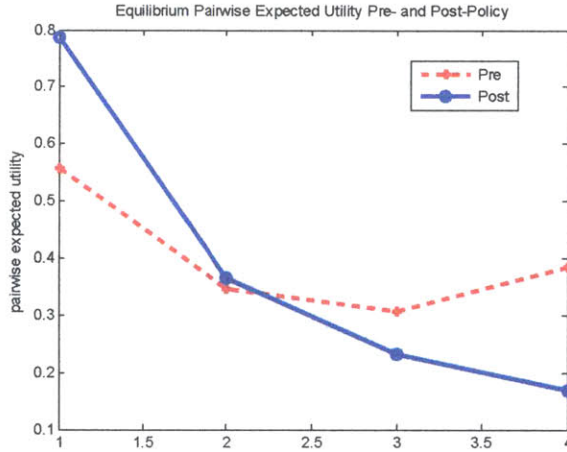
Suppose $G1 = \{1, 2, 3, 4\}$, and $G2 = \{2, 4, 6, 8\}$. I want to focus on the welfare effect of the policy deriving from the endogenous change in equilibrium matching, and shut down the effect deriving from a change in the curvature of $V(p)$. A natural way to do this is to consider a *small but discrete* change in N , from a value of N under which NAM is the unique equilibrium, to a value of N under which PAM is the unique equilibrium.

For example, let $N_1 = 2.1$ and $N_2 = 1.9$. So, the levels of the function change very little, but the marginal variance cost is convex pre-policy and concave post-policy, which we know will shift the unique equilibrium match from NAM to the other extreme.

Equilibrium project choice then experiences the following change: the red line is before the policy, and the blue line is after the policy:



The certainty-equivalent for each matched pair experiences the following change:



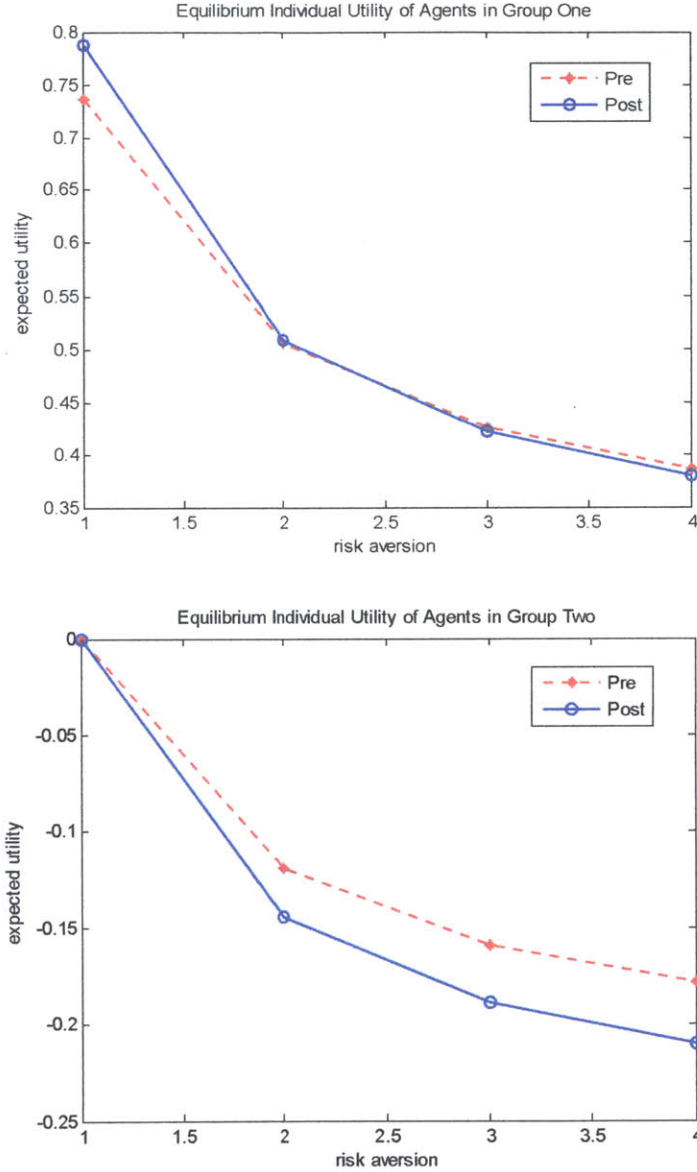
Clearly, this risk-reduction policy is *not* a Pareto improvement. In fact, the policy causes much dispersion and inequality in joint expected utilities post-policy, where the most risk-averse agents partnered with each other are by far the worst off.

However, we might also want to know what happens to individual utilities. Such analysis is challenging in a model of endogenous matching, as the vector of equilibrium individual expected utilities is not unique (please refer to the discussion in the Results section). The natural approach is then to apply relevant selection criteria and focus on the equilibrium vector which survives such selection.

Because risk reduction policies are generally intended to target the most risk-averse people, let us study the vector of equilibrium expected utilities where the most risk-averse agent in group

two²⁸ is best off—that is, the most risk-averse agent in group two has a higher utility in this vector than any other possible vector of equilibrium expected utilities.

Comparing pre- and post-policy individual utilities of group one and group two:



We see that the more risk-averse agents in group one and group two are worse off after implementation of the policy, purely as a result of the endogenous network response: a price stabilization policy which especially reduces the risk of higher mean yield crops, and is intended to reduce the risk burden of the poor, may in fact exacerbate inequality and particularly harm the most risk-averse agents, precisely because the policy causes the least risk-averse agents to abandon their roles

²⁸Equivalently, group one.

as informal insurers of the most risk-averse agents, in favor of entrepreneurial partnerships with fellow less risk-averse agents.

In particular, the least risk-averse agents are better able to manage risk *ex post* when paired with each other rather than with the more risk-averse agents whom they have to insure, and this enables them to take advantage of the decreased aggregate risk and undertake the higher mean projects (e.g. a new technology). On the other hand, the most risk-averse agents are much less able to manage risk *ex post* when paired with each other rather than with their less risk-averse informal insurers. This means that they must rely heavily on *ex ante* methods to manage risk, and so they choose projects with especially low mean returns.

This intuition adds interestingly to a few existing results. Attanasio and Rios-Rull (2000) model the introduction of formal insurance as a policy which reduces the riskiness of the environment. They also find that such a policy may hurt the welfare of the most risk-averse agents. However, their model of (non-endogenous) informal risk-sharing, which builds off Ligon, Thomas, Worrall (2001), centers around limited commitment—two agents interacting repeatedly can only sustain informal risk-sharing with the threat of cutting off all future ties for someone who reneges. This particular type of punishment has the important implication that the value of future ties directly trades off with equilibrium risk-sharing. So, anything that lowers the cost of autarky (remaining alone and being informally uninsured) will decrease the level of informal insurance that can be sustained. They then frame the introduction of formal insurance as a policy that would reduce the cost of being in autarky, and subsequently argue that the level of informal insurance falls, which may reduce aggregate welfare.

However, in this paper, commitment is taken to be perfect. Perfect commitment in Attanasio and Rios-Rull (2000) would preclude the conclusion that the introduction of formal insurance could hurt the most risk-averse agents, because lowering the cost of autarky matters only through the punishment of cutting off future ties, which is no longer relevant. I show that, *even under perfect commitment*, introducing formal insurance might still reduce the welfare of the most risk-averse agents, because the composition of the informal risk-sharing network changes in response. In my model, reducing the riskiness of the environment raises the value of autarky, but it *also* raises the value of being in a relationship, and raises it heterogeneously across partnerships of different risk compositions.

By contrast, Chiappori et al. (2011) estimate that the *least* risk-averse individuals are the ones left worse off after the introduction of formal insurance. The informal argument is that the least

risk-averse agents are displaced as informal insurers. However, this exactly illuminates the need for a model of the *equilibrium network* of relationships—I show that the least risk-averse agents do leave their roles as informal insurers, but only because they prefer to undertake entrepreneurial pursuits instead. It would be interesting to see how the estimation changes after accounting for the endogeneity of matching.

1.6 Testing the Theory: An Experiment

This model provides a natural setting for an experimental test of the theory. In this section, I propose a basic experimental design to test the key predictions. First, I discuss subject and setting choice. Next, I describe important elements of the setup. In particular, I describe how to elicit risk type, and how to design different sets of gamble choices. Finally, I outline the empirical strategy suggested by the testable conditions from Section 4.

1.6.1 Subject and Setting Choice

The theoretical framework is a good fit for an agricultural context. As discussed in the introduction and background sections, farmers in a broad cross-section of regions are demonstrably risk-averse, and heterogeneous in their risk aversion. Moreover, this heterogeneity in risk aversion translates into a heterogeneity in equilibrium behavior; in particular, crop portfolios, spatial plot choices, input and irrigation choices, and openness to innovative techniques vary by farm. Finally, there are plenty of examples of the prevalence of collectivity in farming—farmers seeking advice from each other, farmers entering risk-sharing agreements and self-help groups with each other, and most directly, landowners employing farmers, for example as sharecroppers.

This suggests several advantages in selecting subjects from villages which are primarily agricultural, and, correspondingly, recruiting subjects for whom farming is the primary occupation. This is not because the experimental setup must require subjects to possess farming expertise, but rather because the strong connection between their lives and the task of choosing amongst gambles with differing yield distributions gives us a measure of confidence about the relevance and intuitiveness of the experiment, and their subsequent comprehension of it.

Since matching pattern regressions will be run at the village level, it will be necessary to have a reasonably large sample of villages to obtain power. Because the theoretical result is distribution free, it is able to accommodate heterogeneity in risk type distributions across villages. Villages

should be chosen which do not communicate with each other, to prevent information about the experiment from spreading. Characteristics of the villages should be recorded, such as geographical location, agroclimate, size, primary occupation, average income, and proximity to urban center.

Field officers should hold private interview sessions with randomly-chosen households in each village to extend an invitation to send a member to participate in the experiment. As in Attanasio et al. (2010), this should be done only a few days previous to the experiment, to minimize social selection. Standard characteristics of each interviewed household, and of each adult member of that household, should be collected, such as ethnicity and religion, physical characteristics of the dwelling, occupations, income. If the primary occupation is agricultural, collect data on crops grown, plot size, plot scattering, irrigation, fertilizer, and total harvest.

Additionally, kinship, friendship, neighbor, trust, and employment network data must be collected from households. For example, each interviewed individual may be asked for the names of her relatives, her friends, her neighbors, the people she trusts most, and the people with whom she works.

1.6.2 Setup

Round 1: Risk Elicitation

The experiment will consist of two rounds. In the first round, the "risk type elicitation round", each subject will meet privately with a field researcher, where she will be presented with a set of gambles varying in riskiness and asked to choose one. (The design of this set of gambles is discussed in detail in the subsection below, "Round 2: Joint Project Choice".) The gamble choice will be explained pictorially in a manner that accounts for potential illiteracy and unfamiliarity with probability, and comprehension will be tested. A selection of a gamble with higher mean and variance will indicate a lower degree of risk aversion than a selection of a gamble with lower mean and variance. Additionally, each subject will be asked to report what gamble they think the people they reported trusting the most during the initial household interview would have chosen. This allows us to get a sense of how accurately individuals perceive the risk aversion of those with whom they are close.

After all risk types have been elicited²⁹, subjects will be gathered in the same room, where

²⁹The Rabin (2000) critique does apply here, but many experiments elicit risk in the literature in this way, so by revealed preference the payoff in economic insight seems to be worth it. Moreover, in this paper, if the environment is such that the marginal variance cost function is globally convex or concave, then all we need from the risk elicitation

each individual's gamble choice will be posted publicly, again in a manner that accommodates illiteracy. (Subjects are *not* told initially that their first round gamble choices will be announced.) The purpose of this announcement is to ensure common knowledge of everyone's risk types.

Between Rounds: Matching

Following the elicitation of risk types and the public posting of the gamble choices, subjects shall be randomly divided into two equally-sized groups (if there is an odd number of people, then one person shall remain unmatched). A researcher will then explain that they will be allowed to pair up across groups, and that once paired, each pair will be offered the same set of gambles that they faced in their private first rounds, where they will be allowed to choose a project jointly and commit *ex ante* to a return-contingent sharing rule of the realized income by reporting their agreed-upon rule to a researcher, who will enforce it. (Arguments can be made for or against the possible dependence of the second round gamble choice on first round gamble outcomes if the set of gambles is the same in both rounds, and as Attanasio et al. (2010) have set a precedent of reusing the first round set, I leave discussion of this question to another forum).

After this explanation of the proceedings of the rest of the game, the subjects will be given half an hour to establish partnerships (if an individual chooses to remain unmatched, she will receive 0 for sure because she will not be allowed to select a gamble. Since the supports of all gambles are on the positive real line, nobody should choose to remain unmatched as long as there are partners available).

Round 2: Joint Project Choice

Once partnerships have been established, researchers will visit each pair privately to take down their joint project choice and the sharing rule they have agreed upon. Then, researchers will draw a realization from each pair's project choice, and distribute funds according to the sharing rules that were reported to them. This will mark the end of the experiment.

The key to the experimental test of the theory is the design of the set of gambles from which matched pairs are asked to choose. Recall that the conditions for unique positive-assortative and negative-assortative matching emerging from the model in Section 4 centered on the rate of curvature of the mean-variance trade-off across the spectrum of available projects. In the model, the

round is the *ordinal* risk attitudes.

trade-off between the desire to match with someone with a different risk type in order to smooth a given risk in a mutually satisfying way, and the desire to match with someone with a similar risk type in order to choose a mutually agreeable risk, is captured by this rate of curvature of the mean-variance trade-off. When a slightly higher mean return corresponds with a drastically higher variance of return, so that the difference in variance of a low mean return project and a high mean return project is extremely stark, negative-assortative matching is more likely to result, since all possible pairs prefer the safer, low mean return project. On the other hand, when a small increase in mean return corresponds to a moderately higher increase in variance, positive-assortative matching is more likely to result, since there is significant disagreement about preferred project: less risk-averse people prefer riskier projects, while more risk-averse people prefer safer projects.

Hence, the theory suggests a clear intuition for designing sets of gamble choices: within a given set, gambles should differ in mean and variance of return, with higher mean gambles having correspondingly higher variances, and the shape of this mean-variance trade-off should differ across different sets of gamble choices.

For example, a reasonable approach would be to design two types of gamble sets, each containing five gambles. Villages in the sample would be randomly allocated into two groups, where villages in the first group face the first set of gamble choices, and villages in the second group face the second set of gamble choices. Indexing set by superscript i , and gamble by subscript j , the challenge is to choose values for (p_j^i, V_j^i) , the mean and variance of gamble j in set i , for $i \in \{1, 2\}$ and $j \in \{1, 2, 3, 4, 5\}$.

There are three significant feasibility constraints imposed by the realism of an experiment that prevent us from directly applying the theoretical result to this design problem.

1. The return distributions of the gambles must have *positive, bounded supports*. It seems unwise to extract money from an experimental subject by declaring that she received a "negative realization", and it seems expensive to allow for the possibility of a realization of infinity.

The relevant implication is that we cannot offer gambles with distributions that are literally distributed normally.

2. The set of gamble choices is *discrete*. The key idea in the theoretical model was not that a continuum of projects existed literally, but rather that activity diversification (e.g. intercropping) enabled a spectrum of mean returns with corresponding variances that were higher for higher means (an upper envelope).

In the experiment, subjects are not allowed to create portfolios of projects; instead, they must select a single gamble. This can be thought of as reflecting rigidities in activity diversification (e.g. intercropping is not possible because of soil type), and limits a pair's *ex ante* risk management ability. An individual r 's desire to pair up with another individual r because they share project preferences is dampened by the fact that first, their favorite project is unlikely to be available in the small set of gamble choices, and second, because there is not a wide range of choices, individuals of other risk types will cluster with r on project choice.

The relevant implication of this is that distinct pairs are no longer guaranteed to select distinct projects. Instead, there will be "clustering" on each of the gamble choices; the degree of this "clustering" depends on the distribution of risk types in the population and on the shape of the mean-variance trade-off.

3. *Limited liability* holds in an experimental setting. In the theory, the return-contingent sharing rule of a pair is completely unconstrained. In the experiment, this will not be the case. If for example a gamble's realized return is 0, then each partner must receive 0.

This constraint is the most serious, because it inhibits the transferability of expected utility. For example, if 0 is in the support of a gamble's return distribution, then partners cannot agree to nonzero state-independent transfers. This restriction on the sharing rule limits the pair's *ex post* risk management ability, and, importantly, limits different pairs differently depending on risk type composition. A quite risk-averse individual r 's desire to pair up with a much less risk-averse individual $r' \ll r$ is dampened by the fact that, because of limited liability, r' may not be able to ensure her a return-independent transfer.

Hence, the pairwise sum of certainty-equivalents calculated in Section 2 holds only a bounded range of divisions v of the expected utility pie, where these bounds depend on risk type and therefore influence endogenous partner choice.

Because of these three limitations of the experimental environment, it is necessary to construct an "experimental model" satisfying these constraints which captures the essence of the theory: given a population of heterogeneously risk-averse agents, the risk relationship of the gambles to each other in the available choice set determines whether positive-assortative or negative-assortative matching results.

The idea is the following. Recall that the sum of $N > 2$ independent draws from a $uni[a, b]$ distribution approximates a normal distribution on the support $[Na, Nb]$. For example, if $a = 0$

and $b = 1$, the convolution of the N densities is (figure taken from Grinstead and Snell (1997)):

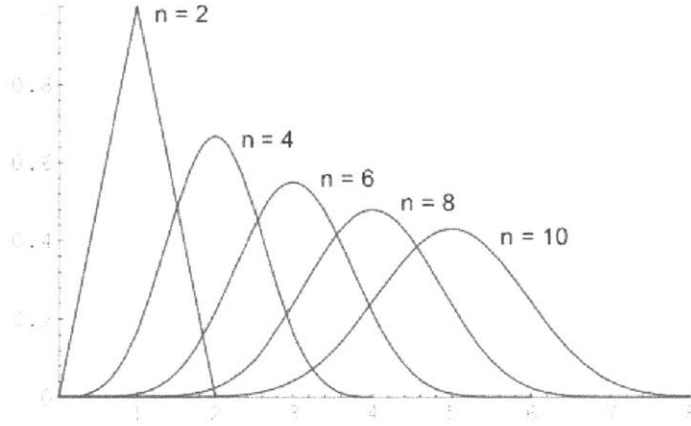


Figure 7.6: Convolution of n uniform densities.

This allows us to offer gambles whose returns are approximately normally-distributed on bounded, positive supports, which enables feasible experimental payments and allows us to apply our understanding of the role played by the mean-variance trade-off from the theoretical results of Section 3.

Therefore, experimental subjects should be asked to choose from one of the following five gambles:

1. Receive the sum of N independent draws from a distribution which yields x for sure. (This is the safe choice.)
2. Receive the sum of N independent draws from a $uni[a_2, b_2]$ distribution, where:

$$\begin{aligned} a_2 &< x < b_2 \\ \frac{a_2 + b_2}{2} &> x \end{aligned}$$

so that this gamble has a higher mean but also a higher variance than the safe choice.

3. Receive the sum of N independent draws from a $uni[a_3, b_3]$ distribution, where:

$$\begin{aligned} a_3 &< a_2 \\ b_3 &> b_2 \\ \frac{a_3 + b_3}{2} &> \frac{a_2 + b_2}{2} \end{aligned}$$

4. Receive the sum of N independent draws from a $uni[a_4, b_4]$ distribution, where:

$$\begin{aligned} a_4 &< a_3 \\ b_4 &> b_3 \\ \frac{a_4 + b_4}{2} &> \frac{a_3 + b_3}{2} \end{aligned}$$

5. Receive the sum of N independent draws from a $uni[a_5, b_5]$ distribution, where:

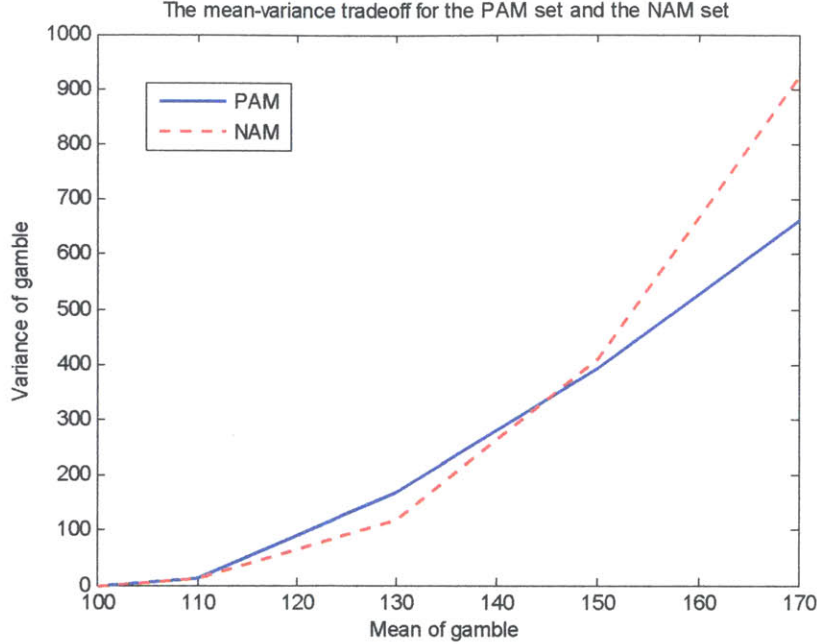
$$\begin{aligned} a_5 &< a_4 \\ b_5 &> b_4 \\ \frac{a_5 + b_5}{2} &> \frac{a_4 + b_4}{2} \end{aligned}$$

Each gamble has an approximately normal distribution on the support $[Na_k, Nb_k]$, and gambles with higher means also have higher variances. The values of a_2, a_3, a_4, a_5 and b_2, b_3, b_4, b_5 must be chosen to shape the curvature of the mean-variance trade-off across gambles so that positive-assortative matching is supported in one set, and negative-assortative matching is supported in the other set.

It is useful to consider a concrete example of how these sets of gamble choices might be designed. Let $N = 10$, so that a matched pair receives the realized sum of 10 independent draws from their chosen gamble. The two sets of gamble choices are:

	PAM	NAM
$x :$	10	10
$[a_2, b_2] :$	[9, 13]	[9, 13]
$[a_3, b_3] :$	[6, 20]	[7, 19]
$[a_4, b_4] :$	[4, 26]	[4, 26]
$[a_5, b_5] :$	[3, 31]	[0, 34]

The following figure shows the mean-variance trade-off across gambles for the PAM set and for the NAM set:



It is worth noting the novelty of the design necessitated by the theory of endogenous risk-sharing pair formation given both *ex ante* and *ex post* risk management. In Attanasio et al. (2010), for example, all experimental subjects faced the same set of gambles, where gambles differed in mean and variance. However, the risk-sharing rule was imposed to be an equal division. Fixing the sharing rule at equal division has a variety of implications: importantly, positive-assortative matching will always result, because the gains-from-trade channel is shut down. Additionally, an equal sharing rule automatically satisfies limited liability. Therefore, any set of gamble choices with gambles varying in riskiness can be chosen for this experiment.

Allowing for *ex post* risk management by endogenizing a risk-sharing group's sharing rule, in addition to the endogenous choice of income stream, entails the precise design of different sets of gamble choices. Since the effect of the trade-off between *ex ante* and *ex post* risk management on equilibrium matching is best distilled with normally-distributed returns, one idea is to replicate such a distribution in a feasible way by summing N independent draws from uniform distributions, which differ in mean and variance.

1.6.3 Empirical Strategy

Recall from Proposition 4 that support for the theory in the data would be observation of at least one of the following:

1. People match positive-assortatively in risk attitudes, and the observed means of equilibrium risky projects are convex in the representative risk tolerance of matched pairs.
2. People match negative-assortatively in risk attitudes, and the observed means of equilibrium risky projects are concave in the representative risk tolerance of matched pairs.

The basic empirical strategy consists of 3 steps. First, given a dataset containing Z villages, index a village by k . Then:

1. For each village k , characterize the equilibrium matching pattern (the correlation between the probability that individual i is matched with individual j and their corresponding risk types r_i and r_j). Is the equilibrium matching pattern the "correct" one, that is, the pattern which was meant to be induced by the set of gamble choices offered to that village?
2. Next, examine the behavior of equilibrium project choices and the risk composition of the matched pairs.
3. Then, characterize the relationship between the equilibrium matching pattern and the equilibrium behavior of project choice in risk types across all villages. If we observe convexity of project choice in the representative risk tolerance of matched pairs corresponding with positive-assortative matching, and/or concavity of project choice in the representative risk tolerance of matched pairs corresponding with negative-assortative matching, then this is support for the theory.

To be a little more specific, first focus on data from a given village k . Suppose that village k was offered the "positive-assortative" set of gamble choices, so we should expect positive-assortative matching.

To glean the matching pattern from the data, set the observation level to be a dyad (i, j) of an individual i in group one matched with an individual j in group two (recall that the experiment participants were randomly divided into two equally-sized groups). If each group has M people in it, then there should be $\frac{M(M-1)}{2}$ total observations.

Define a variable that describes the distance between the risk attitudes of i and j :

$$\Delta_{ij} = |r_i - r_j|$$

Note that this is symmetric: $\Delta_{ij} = \Delta_{ji}$.

Now, define a variable m_{ij} :

$$m_{ij} = \begin{cases} 1, & i \text{ matched with } j \\ 0, & \text{else} \end{cases}$$

So, m_{ij} is also symmetric.

Then, for each village k , regress the linkage variable m_{ij} on the distance between risk attitudes, Δ_{ij} , and the pre-existing networks X_{ij}^k (e.g. kinship):

$$m_{ij}^k = \alpha^k + \mu^k \Delta_{ij}^k + X_{ij}^k + \varepsilon_{ij}^k$$

If $\mu^k > 0$, the matching pattern is closer to being negative-assortative, since large differences in risk attitudes increase the probability that two individuals are linked, while if $\mu^k < 0$, the matching pattern is closer to being positive-assortative, since small differences in risk attitudes increase the probability that two individuals are linked.

Confirm that the matching pattern of village k is the pattern that was meant to be induced by the set of gamble choices offered to village k .

Now, look at matched dyads in village k only, and denote the mean of the equilibrium project choice of a matched (i, j) by p_{ij} .

For each matched dyad (i, j) , calculate the representative risk tolerance:

$$\tilde{H}_{ij} = \frac{1}{r_i} + \frac{1}{r_j}$$

For the matched dyads in each village k , regress:

$$p_{ij}^k = c^k + \theta^k \tilde{H}_{ij}^k + \gamma^k \left(\tilde{H}_{ij}^k \right)^2 + V^k + \eta_{ij}^k$$

where V^k is village fixed effects. (A discussion of other methods for testing convexity may be found in footnote 18 of section 4.1.)

It should be the case that $\theta^k > 0$ for all k , since more risk-tolerant partnerships should choose riskier (higher mean) gambles. A positive $\gamma^k > 0$ indicates that equilibrium project means are convex in risk tolerance, while a negative $\gamma^k < 0$ indicates that equilibrium project means are concave in risk tolerance.

Finally, regress:

$$\mu^k = \kappa + \delta\gamma^k + \nu$$

to ascertain the relationship in village k between the equilibrium matching pattern and the behavior of equilibrium projects in risk types.

Support for the model is $\delta > 0$: this indicates that convexity of mean income in the representative risk tolerance of a linked pair corresponds with positive-assortative matching, while concavity of mean income in the representative risk tolerance of a linked pair corresponds with negative-assortative matching.

1.7 Conclusion

This paper demonstrates the many implications of thinking about informal insurance as the risk-sharing achieved within an equilibrium network of partnerships, as opposed to within a single, isolated partnership, by developing and exploring a theory of endogenous matching between heterogeneously risk-averse agents. I show that the equilibrium matching is driven by the trade-off between two general forces: one, a "gains from trade" force, which favors matching between unlike types, and the other, a "similarity of perspective" force, which favors matching between like types. Specifically, conditional on facing a given risk, a less risk-averse agent and a more risk-averse agent benefit from being matched with each other, because the less risk-averse agent is differentially willing to sell insurance which the more risk-averse agent is differentially willing to buy. But when choosing what risk to face, agents prefer partners with similar risk attitudes. In other words, consumption-smoothing forces push the match to be negative-assortative, while income-smoothing forces push the match to be positive-assortative.

Thus, the equilibrium composition of risk-sharing pairs depends critically on the differential interaction between consumption-smoothing and income-smoothing risk management tools across partnerships of different risk compositions. I show that this differential interaction is cleanly captured by the convexity of the marginal variance cost function in expected returns of projects, and provide testable conditions that tie the theory to data. In particular, I describe observable differences in the environments in which negative- and positive-assortative matching patterns emerge, and use these differences to construct a falsifiability condition based on the dependence of mean returns of equilibrium projects on the risk tolerance of matched pairs. I find preliminary support for the theory in existing literature.

This paper highlights the especial importance of accounting for individual heterogeneity in developing economies—the absence of formal institutions causes individuals to depend on their interactions with each other. Thus, we see that a natural constraint on informal insurance, apart from any imperfections in the contracting environment, or other such standard considerations, is simply that all individuals are risk-averse. However, some individuals are more risk-averse than others, and this enables informal insurance and determines its strength. Furthermore, this equilibrium network approach enables an understanding of how individuals endogenously switch between and assume different informal roles in the economy.

A variety of important policy implications follow from this theory. For example, despite perfect commitment in the contracting environment, a policy which reduces the variance of every available project, and particularly subsidizes higher mean projects, is a strict Pareto improvement when the network of partnerships is assumed to stay fixed, but can be seen to leave many individuals worse off, and be especially harmful for the most risk-averse agents, purely as a result of the endogenous network response: the least risk-averse agents abandon their roles as "informal insurers" of the most risk-averse, preferring entrepreneurial partnerships with fellow less risk-averse agents instead. The least risk-averse agents are better able to manage risk *ex post* when matched with each other instead of with the most risk-averse agents whom they have to insure, and this enables them to take advantage of the risk-reduction policy and undertake higher mean projects (such as a new technology). However, the ability of the most risk-averse agents to manage risk *ex post* worsens when they pair with each other rather than with the less risk-averse informal insurers, and this means they rely heavily on *ex ante* methods to manage risk. As a consequence, they undertake projects with especially low mean returns.

Accounting for endogeneity would also affect the analysis of a variety of land reform policies discussed in the literature. For example, Banerjee (2000) points out that the effects of land redistribution cannot be estimated without first understanding the reasons behind the distribution of land in the status quo. If the distribution of land served a risk-sharing purpose, or alternatively, if agents optimized their risk-sharing activities given the land distribution (for example, large landowners provide informal insurance to the landless), then land redistribution might have no effect, because it is implicitly ignored, or it might actually hurt informal insurance.

Regulation of the rental contract, or the wage, is another reform which should be considered through the lens of this model. We know that employment contracts (such as the share contract) often balance multiple needs, such as insurance provision as well as incentive provision. The results

of this paper show that this contract is a key consumption-smoothing tool—legislation which places restrictions on the set of permissible rental contracts diminishes the power of *ex post* risk management as an insurance tool, and pushes the equilibrium matching towards positive-assortativity.

Other reforms affect the adaptability of the network. For example, Banerjee et al. (2002) discuss Operation Barga, a 1970s West Bengali land reform which reduced the ability of landlords to expel their tenants. An unexplored effect of such a reform is the added matching friction: relationships will be unable to respond optimally to changes in the environment. For example, if a bad rainfall shock unexpectedly hit a farming village, increasing the steepness of the mean-variance trade-off across crop choices, then in the absence of matching frictions, less risk-averse individuals would act as informal insurers of the more risk-averse. But Barga may have locked less risk-averse individuals into their relationships with other less risk-averse individuals.

Finally, accounting for endogeneity may substantially alter our inferences from empirical observations. For example, an observation that the strength of dependence of a sharecropper's rent contract on realized harvest is *positively* correlated with the riskiness of the crop grown may tempt us to conclude that risk is not a problem for farmers, since risk considerations do not appear to influence contract design (Allen and Lueck (1992)). However, if more risk-averse farmers work for less risk-averse landowners, cultivating safer crops under low-powered contracts, while less risk-averse farmers work for more risk-averse landowners, cultivating riskier crops under high-powered contracts, the same empirical observations would emerge, but risk concerns would be playing a significant role in contract design, through the unaccounted-for channel of contracting partner choice.

All sorts of challenges remain. This model can clearly be enriched in many ways—by generalizing preferences and the project space, by introducing limited commitment or incomplete information, by allowing for more complex forms of group formation, or by adding dynamics to study the feedback between riskiness of project choice and investment capability, to characterize a potential "innovation trap". The main contribution of this paper is to propose a tractable model of the endogenous formation of risk-sharing relationships, upon which further dimensions can be layered, and to show that even this simple model can lead to substantial changes in our understanding of numerous policy debates. The unique environments of developing economies make design problems particularly challenging, but it is the magnitude of the stakes which makes them particularly important.

1.A Appendix

1.A.1 The model with general distribution of returns

Suppose all elements of the environment described in Section 3 are the same, except the returns of a project p are now assumed to follow a more general distribution: $R_p \sim f(R_p|p)$, where the support $\Omega(R_p) = (-\infty, \infty)$. Assume that $f(R_p|p)$ is a well-defined pdf whose cumulants exist (recall that the cumulant-generating function is the log of the moment-generating function).

Given that the pair (r_1, r_2) is matched, agent r_1 who must ensure her partner some level of expected utility $-e^{-v}$ solves the following problem:

$$\begin{aligned} \max_{s(R_p)} \int_{-\infty}^{\infty} -e^{-r_1[R_p - s(R_p)]} f(R_p|p) dR_p \text{ s.t.} \\ \int_{-\infty}^{\infty} -e^{-r_2 s(R_p)} f(R_p|p) dR_p \geq -e^{-v} \end{aligned}$$

Then the optimal sharing rule is still linear:

$$s^*(R_p) = \frac{r_1}{r_1 + r_2} R_p + \frac{1}{r_2} \log \int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p) dR_p + \frac{v}{r_2}$$

where the optimal project $p^*(r_1, r_2)$ maximizes the expected utility of r_1 given equilibrium sharing rule $s^*(R_p; v)$:

$$\max_p -e^{\frac{r_1}{r_2} v} \left[\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p) dR_p \right]^{1 + \frac{r_1}{r_2}}$$

Then the sum of certainty-equivalents of the pair is:

$$CE(r_1, r_2) = - \left(\frac{1}{r_1} + \frac{1}{r_2} \right) \log \int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p^*(r_1, r_2)) dR_p$$

Then the representative risk tolerance of the matched pair is:

$$\begin{aligned} \tilde{H}(r_1, r_2) &= \frac{1}{r_1} + \frac{1}{r_2} \\ &= \frac{r_1 + r_2}{r_1 r_2} \end{aligned}$$

and the representative risk aversion of the matched pair is the reciprocal:

$$\hat{H}(r_1, r_2) = \frac{r_1 r_2}{r_1 + r_2}$$

We know that a sufficient condition for the unique equilibrium matching pattern to be PAM is $\frac{d^2 CE(r_1, r_2)}{dr_1 dr_2} > 0$, while we know that $\frac{d^2 CE(r_1, r_2)}{dr_1 dr_2} < 0$ is sufficient for NAM to be the unique equilibrium matching pattern. What conditions are needed for these two inequalities to hold?

Observe that the sum of certainty-equivalents is actually just:

$$\begin{aligned} CE(r_1, r_2) &= -\frac{1}{\hat{H}(r_1, r_2)} C_{R_p}(t = -\hat{H}(r_1, r_2)) \\ &= -\hat{H}(r_1, r_2)p + \sum_{n=1}^{\infty} \frac{\hat{H}(r_1, r_2)^{2n}}{(2n)!} k_{2n}(p) \end{aligned}$$

where $C_{R_p}(t)$ is the cumulant-generating function of the random variable R_p . The notation $k_{2n}(p)$ denotes the $2n^{th}$ order cumulant.

At this point, I add the additional assumption of a *symmetric* distribution, since we know that all higher-order odd cumulants of symmetric distributions are zero. Then it is clear that a given risk-averse agent r evaluates a project p by balancing the benefit of the expected return with the cost, captured by the sum of all the higher order cumulants. The sum of all higher order cumulants can thus be thought of as a "generalized variance" for the purposes of this model, so that the spectrum of risky projects is parameterized by a "generalized mean-variance trade-off".

Lemma 9 *Under a general symmetric distribution, the optimal project choice of a pair minimizes the cumulant-generating function evaluated at negative the representative risk aversion, $-\hat{H}(r_1, r_2)$.*

The first-order condition characterizing optimal project choice is:

$$FOC_p : \sum_{n=1}^{\infty} \frac{\hat{H}(r_1, r_2)^{2n}}{(2n)!} k'_{2n}(p) = \hat{H}(r_1, r_2)$$

We need the following condition for convexity:

$$SOC_p : \sum_{n=1}^{\infty} \frac{\hat{H}(r_1, r_2)^{2n}}{(2n)!} k''_{2n}(p) \geq 0$$

Then differentiate the first-order condition implicitly with respect to \hat{H} :

$$\frac{dp}{d\hat{H}} = \frac{1 - \sum_{n=1}^{\infty} \frac{\hat{H}(r_1, r_2)^{2n-1}}{(2n-1)!} k'_{2n}(p)}{\sum_{n=1}^{\infty} \frac{\hat{H}(r_1, r_2)^{2n}}{(2n)!} k''_{2n}(p)}$$

But the FOC tells us:

$$\begin{aligned} \sum_{n=1}^{\infty} \frac{\hat{H}(r_1, r_2)^{2n}}{(2n)!} k'_{2n}(p) &= \hat{H}(r_1, r_2) \Rightarrow \\ \sum_{n=1}^{\infty} \frac{\hat{H}(r_1, r_2)^{2n-1}}{(2n)!} k'_{2n}(p) &= 1 \end{aligned}$$

and

$$\sum_{n=1}^{\infty} \frac{\hat{H}(r_1, r_2)^{2n-1}}{(2n-1)!} k'_{2n}(p) > \sum_{n=1}^{\infty} \frac{\hat{H}(r_1, r_2)^{2n-1}}{(2n)!} k'_{2n}(p) = 1$$

Then:

$$\frac{dp}{d\hat{H}} < 0$$

This proves the following lemma.

Lemma 10 *The optimal project choice for a matched pair in this generalized model is decreasing in representative risk aversion, and furthermore, $p^*(r_1, r_2)$ is symmetric in r_1, r_2 since $\hat{H}(r_1, r_2)$ symmetric in r_1, r_2 .*

Hence, in this generalized model, the condition on the fundamentals of the model for the equilibrium matching is still a parameteric condition on the generalized mean-variance trade-off, which can be calculated.

Specifically, we know:

$$CE(\tilde{H}(r_1, r_2)) = -\frac{1}{\tilde{H}(r_1, r_2)} p\left(\tilde{H}(r_1, r_2)\right) + \sum_{n=1}^{\infty} \frac{1}{(2n)! \tilde{H}(r_1, r_2)^{2n}} k_{2n}\left(p\left(\tilde{H}(r_1, r_2)\right)\right)$$

and the optimal project choice of a matched pair with risk tolerance \tilde{H} is:

$$\sum_{n=1}^{\infty} \frac{1}{(2n)! \tilde{H}(r_1, r_2)^{2n}} k'_{2n}\left(p\left(\tilde{H}(r_1, r_2)\right)\right) = \frac{1}{\tilde{H}(r_1, r_2)}$$

Lemma 11 *A sufficient condition for unique PAM is the convexity of $CE(\tilde{H})$ in \tilde{H} , while a sufficient condition for unique NAM is the concavity of $CE(\tilde{H})$ in \tilde{H} .*

1.A.2 The model nests a framework with individual project choice, pooled income, and individual rationality

Suppose that, as in the benchmark model described in Section 3, a spectrum of projects $p \geq 0$ is available, where $R_p \sim N(p, V(p))$, and $V(p) > 0$, $V'(p) > 0$, and $V''(p) > 0$. However, instead of jointly choosing a project, each agent r in the population chooses a project p_r , and a matched pair pools their realized incomes and splits it according to a pre-decided sharing rule contingent on pooled realized returns, $s(R_{p_{r_1}} + R_{p_{r_2}})$.

Assume that the income realizations of r_1 and r_2 follow distributions $f(R_{p_{r_1}}|p_{r_1})$ and $f(R_{p_{r_2}}|p_{r_2})$, respectively (partners' incomes may be correlated in some way). Thus, the distribution of pooled income, $f(R_{p_{r_1}} + R_{p_{r_2}}|p_{r_1}, p_{r_2})$ can be calculated from the individual distributions and the structure of the correlation (for example, if incomes were independent, the distribution of pooled income would simply be the convolution of the individual distributions).

Observe that IR is clearly satisfied, because the autarkic sharing rule (the sharing rule which specifies that each person keeps her own realized income) is always an option for the pair. Hence, all individuals prefer to match over remaining unmatched.

But now it is clear that matched individuals in this framework are essentially jointly choosing a distribution of pooled income. Thus, since IR is always satisfied, the framework where individuals do not remain unmatched and partnerships jointly choose projects is equivalent. Moreover, this latter framework illustrates much more cleanly the dependence of the equilibrium matching on the trade-off between consumption-smoothing and income-smoothing. Hence, this is the model I work with in this paper.

1.A.3 Proof that NTU problem has TU representation in expected utility

We know from Appendix 1 (which discusses the generalized model) that the indirect utility of an agent r_1 who ensures his partner, r_2 , a level of expected utility $-e^{-v}$ is:

$$\phi(r_1, r_2, v) = -e^{\frac{r_1}{r_2}v} \left[\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p) dR_p \right]^{1 + \frac{r_1}{r_2}}$$

Partner r_2 receives:

$$\begin{aligned}
EU_{r_2}(v) &= -e^{-r_2 \frac{v}{r_2}} \\
&= -e^{-v}
\end{aligned}$$

Then the certainty-equivalent of each member is:

$$\begin{aligned}
CE_{r_1}(v) &= -\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \log \left[\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p | p^*(r_1, r_2)) dR_p \right] - \frac{v}{r_2} \\
CE_{r_2}(v) &= \frac{v}{r_2}
\end{aligned}$$

There is clearly a one-to-one trade-off between the certainty-equivalent of r_1 and the certainty-equivalent of r_2 . In other words, the slope of the Pareto frontier of expected utility (modulo a monotonic transformation) is -1 . Hence, expected utility is transferable.

It is also interesting to see that the generalized increasing/decreasing differences technique from Legros and Newman (2007) can also be used directly to prove the main matching result of this model. (We know that generalized increasing/decreasing differences is the same as supermodularity/submodularity in a transferable utility setting; it is simply of interest to see that the Legros and Newman technique can be directly applied to this model of risk-sharing which allows for endogenous choice of risk.)

Recall the Legros and Newman generalized differences condition for matching under nontransferable utility: the following is a sufficient condition for positive-assortative matching to be the unique equilibrium matching pattern:

$$\begin{aligned}
\phi(r_1, r_2, p^*(r_1, r_2), v) &= \phi(r_1, r'_2, p^*(r_1, r'_2), u) \Rightarrow \\
\phi(r'_1, r_2, p^*(r'_1, r_2), v) &< \phi(r'_1, r'_2, p^*(r'_1, r'_2), u) \quad \forall r_1, r_2
\end{aligned}$$

We want to show that this condition corresponds to the condition that $\frac{d^2 CE(r_1, r_2)}{dr_1 dr_2} > 0$, where $CE(r_1, r_2)$ is the sum of certainty-equivalents in a matched pair (r_1, r_2) .

We know from the derivation of the optimal sharing rule and the sum of certainty-equivalents in the general case that the indirect utility of agent r_1 when matched with agent r_2 and ensuring him expected utility $-e^{-v}$ is:

$$\phi(r_1, r_2, p^*(r_1, r_2), v) = -e^{\frac{r_1}{r_2}v} \left[\int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p | p^*(r_1, r_2)) dR_p \right]^{1 + \frac{r_1}{r_2}}$$

So

$$\phi(r_1, r_2, p^*(r_1, r_2), v) = \phi(r_1, r'_2, p^*(r_1, r'_2), u) \Leftrightarrow$$

$$-e^{\frac{r_1}{r_2}v} \left[\int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p | p^*(r_1, r_2)) dR_p \right]^{1 + \frac{r_1}{r_2}} = -e^{\frac{r_1}{r'_2}u} \left[\int e^{-\frac{r_1 r'_2}{r_1 + r'_2} R_p} f(R_p | p^*(r_1, r'_2)) dR_p \right]^{1 + \frac{r_1}{r'_2}} \Leftrightarrow$$

$$\begin{aligned} \frac{r_1}{r_2}v + \left(1 + \frac{r_1}{r_2}\right) \log \int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p | p^*(r_1, r_2)) dR_p &= \\ \frac{r_1}{r'_2}u + \left(1 + \frac{r_1}{r'_2}\right) \log \int e^{-\frac{r_1 r'_2}{r_1 + r'_2} R_p} f(R_p | p^*(r_1, r'_2)) dR_p &\Leftrightarrow \end{aligned}$$

$$\begin{aligned} \frac{1}{r'_2}u - \frac{1}{r_2}v &= \left(\frac{1}{r_1} + \frac{1}{r_2}\right) \log \int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p | p^*(r_1, r_2)) dR_p - \\ &\quad \left(\frac{1}{r_1} + \frac{1}{r'_2}\right) \log \int e^{-\frac{r_1 r'_2}{r_1 + r'_2} R_p} f(R_p | p^*(r_1, r'_2)) dR_p \end{aligned}$$

Then the condition $\phi(r'_1, r_2, p^*(r'_1, r_2), v) < \phi(r'_1, r'_2, p^*(r'_1, r'_2), u)$ is equivalent to:

$$\begin{aligned} &\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \log \int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p | p^*(r_1, r_2)) dR_p - \\ &\left(\frac{1}{r_1} + \frac{1}{r'_2}\right) \log \int e^{-\frac{r_1 r'_2}{r_1 + r'_2} R_p} f(R_p | p^*(r_1, r'_2)) dR_p > \\ &> \left(\frac{1}{r'_1} + \frac{1}{r_2}\right) \log \int e^{-\frac{r'_1 r_2}{r'_1 + r_2} R_p} f(R_p | p^*(r'_1, r_2)) dR_p - \\ &\left(\frac{1}{r'_1} + \frac{1}{r'_2}\right) \log \int e^{-\frac{r'_1 r'_2}{r'_1 + r'_2} R_p} f(R_p | p^*(r'_1, r'_2)) dR_p \end{aligned}$$

That is, the condition is equivalent to increasing differences in r_1 and r_2 of the following ex-

pression:

$$-\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \log \int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p | p^*(r_1, r_2)) dR_p$$

Now, let's find the certainty-equivalent of r_1 when she is paired with r_2 and ensuring him $-e^{-v}$:

$$\begin{aligned} -e^{-r_1 CE_{r_1}} &= \phi(r_1, r_2, p^*(r_1, r_2), v) \Rightarrow \\ CE_{r_1} &= -\frac{1}{r_1} \log [-\phi(r_1, r_2, p^*(r_1, r_2), v)] \end{aligned}$$

And the certainty-equivalent of r_2 is $\frac{v}{r_2}$. Then the sum $CE(r_1, r_2)$ is:

$$\begin{aligned} CE(r_1, r_2) &= -\frac{1}{r_1} \log [-\phi(r_1, r_2, p^*(r_1, r_2), v)] + \frac{v}{r_2} \\ &= -\frac{1}{r_1} \left[\frac{r_1}{r_2} v + \left(1 + \frac{r_1}{r_2}\right) \log \int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p | p^*(r_1, r_2)) dR_p \right] + \frac{v}{r_2} \\ &= -\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \log \int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p | p^*(r_1, r_2)) dR_p \end{aligned}$$

But this is exactly the same condition we got from the generalized differences condition from Legros-Newman.

The argument is analogous for showing that the Legros and Newman generalized decreasing differences condition for negative-assortative matching to be the unique equilibrium matching pattern corresponds to the condition that $\frac{dCE(r_1, r_2)}{dr_1 dr_2} < 0$.

1.A.4 Proof of main matching result

The cumulant-generating function for a normal distribution $N(p, V(p))$ has a very nice form:

$$C_{R_p}(t = -\hat{H} | p) = -\hat{H}p + \frac{\hat{H}^2}{2} V(p)$$

(See Appendix 1 for the generalized case, which illustrates the role played by the cumulant-generating function in the optimal project choice problem of a partnership.)

Hence the sum of certainty-equivalents in a partnership with representative risk aversion \hat{H} is:

$$CE(\hat{H}) = p - \frac{1}{2}\hat{H}V(p)$$

Recall that:

$$\hat{H} = \frac{r_1 r_2}{r_1 + r_2}$$

Then:

$$\frac{d^2 CE(r_1, r_2)}{dr_1 dr_2} = -\frac{r_1 r_2}{(r_1 + r_2)^3} V\left(V'^{-1}\left(\frac{2(r_1 + r_2)}{r_1 r_2}\right)\right) + \frac{2}{r_1 r_2 (r_1 + r_2)} V'^{-1'}\left(\frac{2(r_1 + r_2)}{r_1 r_2}\right)$$

where the project chosen optimally by a pair (r_1, r_2) is:

$$p^*(r_1, r_2) = V'^{-1}\left[\frac{2(r_1 + r_2)}{r_1 r_2}\right]$$

Therefore, $\frac{d^2 CE(r_1, r_2)}{dr_1 dr_2} > 0$ iff:

$$\frac{2(r_1 + r_2)^2}{r_1^2 r_2^2} V'^{-1'}\left(\frac{2(r_1 + r_2)}{r_1 r_2}\right) > V\left(V'^{-1}\left(\frac{2(r_1 + r_2)}{r_1 r_2}\right)\right)$$

Then, define:

$$g(x) = \frac{x^2}{2} V'^{-1'}(x) - V(V'^{-1}(x))$$

Finding a sufficient condition for PAM is equivalent to finding conditions on $V(\cdot)$ such that $g(x) > 0$ for $x > 0$.

Assume $g(0) = 0$, that is, $V(0) = 0$ and $V'(0) = 0$. Then:

$$g'(x) = \frac{x^2}{2} V'^{-1''}(x)$$

So, $g'(x) > 0$ if $V'^{-1''}(x) > 0 \Leftrightarrow V'''(x) < 0$ for $x > 0$, and $g'(x) < 0$ if $V'^{-1''}(x) < 0 \Leftrightarrow V'''(x) > 0$, which proves the result.

On the Corollaries: recall that optimal project choice of (r_1, r_2) is given by:

$$p^*(r_1, r_2) = V'^{-1} \left[\frac{2(r_1 + r_2)}{r_1 r_2} \right]$$

The sufficient condition for PAM to be the unique equilibrium is $V'^{-1''}(p) > 0$ ($V'''(p) < 0$). It is clear that this implies project choice will be convex in the representative risk tolerance of the pair, $\frac{2(r_1+r_2)}{r_1 r_2} = \frac{1}{r_1} + \frac{1}{r_2}$. Similarly, if $V'^{-1''}(p) < 0$, so that NAM is the unique equilibrium, it is clear that project choice will be concave in the representative risk tolerance of the pair.

How does project choice vary in individual type?

$$\begin{aligned} \frac{dp}{dr_1} &= -V'^{-1'} \left[\frac{2(r_1 + r_2)}{r_1 r_2} \right] \left(\frac{2}{r_1^2} \right) \\ \frac{d^2 p}{dr_2 dr_1} &= V'^{-1''} \left[\frac{2(r_1 + r_2)}{r_1 r_2} \right] \frac{4}{r_1^2 r_2^2} \end{aligned}$$

Clearly, $\frac{d^2 p}{dr_2 dr_1} > 0$ iff $V'^{-1''}(p) > 0$, and $\frac{d^2 p}{dr_2 dr_1} < 0$ iff $V'^{-1''}(p) < 0$.

Now, let's look more closely at the Sharpe ratio of a project p when $V(p) = p^N$, $N > 1$:

$$\begin{aligned} SR(p) &= \frac{p}{V(p)^{\frac{1}{2}}} \\ &= p^{1-\frac{N}{2}} \end{aligned}$$

Then:

$$\frac{dSR(p)}{dp} = \left(1 - \frac{N}{2} \right) p^{-\frac{N}{2}}$$

Then

$$\begin{aligned} \frac{dSR(p)}{dp} &> 0 \text{ if } N \in (1, 2) \\ &< 0 \text{ if } N > 2 \end{aligned}$$

But $N \in (1, 2)$ is precisely the condition for PAM to be the unique matching equilibrium, and $N > 2$ is precisely the condition for NAM to be the unique matching equilibrium.

1.A.5 Proof that NAM is the unique equilibrium match when income-smoothing is shut down

Suppose that all pairs must undertake the same project, p . For instance, the government mandates that all farmers must grow rice. This effectively shuts down the income-smoothing channel.

Differentiate $CE(r_1, r_2)$ with respect to r_1 and r_2 when there is no project choice, so that all pairs (r_1, r_2) face the same risky income stream $f(R_p|p)$. The cross-partial $\frac{d^2 CE(r_1, r_2)}{dr_1 dr_2}$ is:

$$\frac{r_1 r_2}{(r_1 + r_2)^3} \frac{\int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p) dR_p \int R_p^2 e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p) dR_p - \left[\int R_p e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p) dR_p \right]^2}{\left[\int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p) dR_p \right]^2}$$

But we know that:

$$\int e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p) dR_p \int R_p^2 e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p) dR_p > \left[\int R_p e^{-\frac{r_1 r_2}{r_1 + r_2} R_p} f(R_p|p) dR_p \right]^2$$

since we know variance is always positive. Therefore:

$$\int f(R_p|p) dR_p \int R_p^2 f(R_p|p) dR_p > \left[\int R_p f(R_p|p) dR_p \right]^2$$

and $g(R_p) = e^{-\frac{r_1 r_2}{r_1 + r_2} R_p}$ is a convex function.

Hence:

$$\frac{d^2 CE(r_1, r_2)}{dr_1 dr_2} < 0$$

and negative-assortative matching therefore results as the unique equilibrium.

This corresponds with the result from Chiappori and Reny (2006) and Schulhofer-Wohl (2006).

1.A.6 Proof that PAM is the unique equilibrium match when consumption-smoothing is shut down

We know from the main matching result that the optimal sharing rule is linear. Suppose the slope of the sharing rule $s(R_p) = a + bR_p$ is fixed at b for all possible pairs of risk types. (For example, the government mandates an equal split of the output, but partners are free to make fixed transfers

to each other.)

This removes consumption-smoothing as an effective risk management tool, leaving only income-smoothing (project choice). What happens to equilibrium risk-sharing relationships?

A matched pair (r_1, r_2) chooses the relationship-specific transfer a and project p :

$$\max_{a,p} E[-e^{-r_1[R_p - a - bR_p]} | p] \text{ s.t.}$$

$$E[-e^{-r_2[a + bR_p]} | p] \geq -e^{-v}$$

The fixed transfer just serves to satisfy the constraint, since the division of output is fixed at b by the government. Therefore:

$$a = \frac{v}{r_2} + \frac{r_2}{2} b^2 V(p) - bp$$

Then, the project chosen is:

$$\max_p -e^{-r_1 \left[p - \frac{v}{r_2} - \left(\frac{r_1}{2} (1-b)^2 + \frac{r_2}{2} b^2 \right) V(p) \right]}$$

$$p^*(r_1, r_2) = V'^{-1} \left(\frac{2}{r_1(1-b)^2 + r_2 b^2} \right)$$

Then the sum of certainty-equivalents of a given pair (r_1, r_2) is:

$$CE(r_1, r_2) = V'^{-1} \left(\frac{2}{r_1(1-b)^2 + r_2 b^2} \right) - \left(\frac{r_1}{2} (1-b)^2 + \frac{r_2}{2} b^2 \right) V \left(V'^{-1} \left(\frac{2}{r_1(1-b)^2 + r_2 b^2} \right) \right)$$

And:

$$\frac{dCE}{dr_1} = -\frac{(1-b)^2}{2} V \left(V'^{-1} \left(\frac{2}{r_1(1-b)^2 + r_2 b^2} \right) \right)$$

$$\frac{d^2 CE}{dr_2 dr_1} = V'^{-1} \left(\frac{2}{r_1(1-b)^2 + r_2 b^2} \right) \frac{2b^2(1-b)^2}{(r_1(1-b)^2 + r_2 b^2)^3} > 0$$

Hence, positive-assortative matching arises as the unique equilibrium.

So, if the government regulates wages by fixing the slope of the sharing rule at some $b \in [0, 1]$,

where pairs can still make within-pair state-independent transfers, the unique equilibrium matching pattern is always positive-assortative, verifying our intuition that, because consumption-smoothing is held fixed, the "similarity of decisionmaking framework" dominates and people match with people who are like them because they will agree about project choice. This can be thought of as the counterpoint to holding income-smoothing fixed (Appendix 5), as in Chiappori and Reny (2006) and Schulhofer-Wohl (2006)—a policy example suppressing choice of income stream might be, say, the government specifying that only wheat can be grown.

What are some implications of this understanding? The government may be motivated by equality concerns to specify an equal division of output in every relationship, but this may actually generate even more inequality by weakening the informal risk-management toolkit available to individuals, which then triggers endogenous change in risk-sharing networks. Specifically, if agents had been matched negative-assortatively in the status quo (because the "cost function" of project mean is quite convex, say), then this imposition of wage equality leads to positive-assortative matching, which may actually exacerbate inequality: there is a bigger spread in projects, with less risk-averse agents on projects with much higher expected returns while more risk-averse agents are on projects with much smaller expected returns, and less risk-averse agents abandon their roles as informal insurers, and more risk-averse agents wind up bearing more risk.

Finally, note that although this proof is for the case of linear rules with fixed slope but flexible transfer, the intuition is robust to more flexible ways of restricting consumption-smoothing via regulation of the sharing rule. The linearity has nothing to do with the intuition—rather, it allows consideration of a case where consumption-smoothing is held fixed, but transfers can still be made so that the outside options of individuals are still endogenously-determined. (If the transfer were held fixed as well, then utility would be completely nontransferable, and we would be able to calculate each individual's expected utility from being partnered with another individual, without any consideration of endogeneous *vs.* But this is not an interesting case to study.)

1.A.7 Differently-sized groups

We know by now (see, e.g., Appendix 4) that the sum of certainty-equivalents for a given pair (r_1, r_2) in the benchmark case is given by:

$$CE(r_1, r_2) = V'^{-1} \left[\frac{2(r_1 + r_2)}{r_1 r_2} \right] - \frac{r_1 r_2}{2(r_1 + r_2)} V \left[V'^{-1} \left[\frac{2(r_1 + r_2)}{r_1 r_2} \right] \right]$$

Then:

$$\begin{aligned}\frac{dCE}{dr_2} &= -\frac{r_1^2}{2(r_1 + r_2)^2} V \left[V'^{-1} \left[\frac{2(r_1 + r_2)}{r_1 r_2} \right] \right] \\ &< 0\end{aligned}$$

Hence, $CE(r_1, r_2)$ is decreasing in risk aversion.

Thus, if $|G1| < |G2|$, for example, it is the most risk-averse excess agents of $G2$ who will remain unmatched in equilibrium. The main matching results apply to the rest of the agents.

1.A.8 Background and empirical context

The purpose of this section is to provide further empirical context for the model. First, I will discuss the substantial role played by informal insurance motivations in building relationships in risky environments with missing formal insurance and credit markets. Additionally, I will show that risk attitudes are a significant determinant of risk-sharing partner choice.

Next, I will provide evidence that there is a great deal of heterogeneity in risk aversion across individuals in a wide range of settings.

Finally, I will provide evidence of heterogeneity in the riskiness of activities available to individuals, as well as heterogeneity in the relative riskiness of these activities across different environments.

To fix ideas, it may be helpful to envision an agricultural setting, which captures nicely the key elements of the model. Much of the literature discussed in this section is drawn from an agricultural context, where landowners and farmers are heterogeneous in the extent of their risk aversion, and landowners must decide which farmers to work with. Different crops have different yield and profit distributions: some crops are very robust to drought but correspondingly tend to produce low yields on average ("safe" crops), while other crops have the potential for very high yields, but are extremely sensitive to rainfall and other inputs, and blight easily ("risky" crops). In addition to crop portfolio and plot locations, fertilizer and other inputs, irrigation, planting times, and general farming methods and technologies must also be chosen.

Furthermore, the yield and profit distribution of each crop varies across agroclimatic region. Different parts of the world experience different levels of rainfall, soil quality, irrigation, elevation, heat, and other such ecological characteristics, and this influences the stochastic yield and profit of each crop. It is no surprise, then, that equilibrium cropping methods, crop mixes, and contracting

institutions vary so widely across region. A goal of this paper is to advance the understanding of these differences.

Risk Attitude and Informal Insurance Relationships

An abundance of work discusses the considerable role of informal insurance concerns in network formation. People rely on each other to smooth consumption risk and income risk in a wide variety of ways (Alderman and Paxson (1992), Morduch (1995)). A very prevalent consumption-smoothing technique between people is transfers and remittances, and much work has been done to study the nature of the transfers that can be sustained given a risk-sharing group, the shapes of equilibrium networks holding fixed some transfer rule, and who is empirically observed to make transfers with whom (Townsend (1994), Fafchamps and Lund (2003), Genicot and Ray (2003), Bramoulle and Kranton (2007), and Ambrus et al. (2013), to name a few). A general message these papers convey is that the need to manage risk in the absence of formal insurance institutions has huge effects on interpersonal relationships among the poor.

In fact, risk management can affect relationship formation in very specific ways. Rosenzweig and Stark (1989) show that daughters are often strategically married to villages located in environmentally dissimilar regions with minimally correlated farming incomes, for the purposes of consumption-smoothing; households exposed to more income risk are more likely to invest in longer-distance marriage arrangements. Ligon et al. [cite], Fafchamps [1999], and Kocherlakota [1996], among many others, analyze a pure risk-sharing relationship between two heterogeneously risk-averse households who perfectly observe each other's income. Akerberg and Botticini [cite] study agricultural contracting in medieval Tuscany, and find evidence that heterogeneously risk-averse tenant farmers and landlords strategically formed sharecropping relationships based on differing risk attitudes, motivated by risk management concerns. Hence, informal insurance motivations play a substantial role in the formation of relationships.

But how much do individuals care about the risk attitudes of potential partners when forming risk-sharing groups? Naturally, there are many other reasons people might match with each other, but the point of the model is to focus on one important determinant of risk-sharing relationship formation, and to study how equilibrium matching patterns shift along that dimension. Furthermore, there is a great deal of evidence that the risk attitudes of partners are indeed a primary determinant of risk-sharing partner choice. Akerberg and Botticini (2002) provide empirical evidence supporting the presence of endogeneity of matching along risk attitude of landowners and

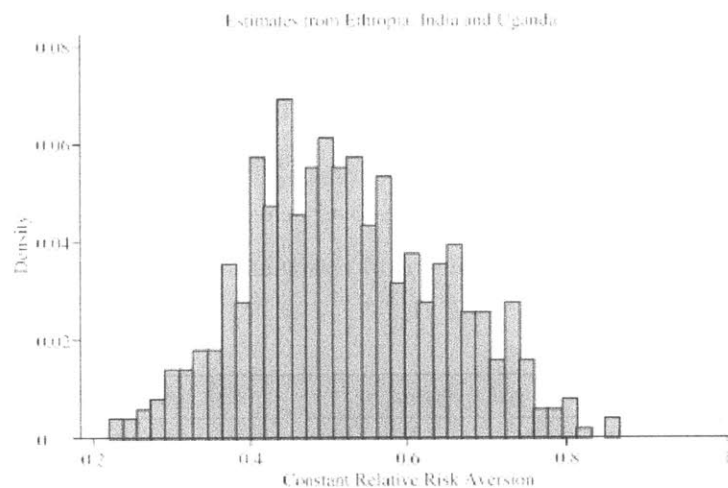
sharecroppers in medieval Tuscany. In their data, they find that share contracts were associated with the safer crop of cereal, while fixed rent (residual claimancy) contracts were associated with the riskier crop of vines. They argue that this is the outcome of endogenous matching—risk-neutral tenants may have been assigned to the riskier crops, resulting in fixed rent contracts for vines, while risk-averse tenants may have been assigned to the safer crops, resulting in share contracts on cereals.

Additional evidence for the importance of risk attitudes as a determinant of risk-sharing relationships is found in Gine et al. (2010) and Attanasio et al. (2012). Gine et al. (2010) run an experiment on small-scale entrepreneurs in urban Peru and allow joint liability groups to form endogenously in a microfinance setting. They find strong evidence of assortative matching along risk attitude. Attanasio et al. (2012) run a unique experiment with 70 Colombian communities. They gather data about risk attitudes and pre-existing kinship/friendship networks, and then allow individuals to form risk-sharing groups of any size. Attanasio et al. find that, when members know each other's risk types, and trust each other (family and friends are in the same group), conditioning on all other potential reasons for matching which they are able to account for (gender, age, geography), there is strong evidence of positive assortative matching along risk aversion.

To further emphasize the significant role of risk attitude in determining risk-sharing relationship formation, I use the dataset from Attanasio et al. (2010) to calculate the proportion of formed links that involved at least one family or one friendship tie, for each municipality. The mean of these proportions is 0.005, or 0.5%. Since it's possible that there were very few family and friend ties reported in the entire dataset to begin with, I also calculate the proportion of all possible links that could have involved at least one family or friendship tie, for each municipality. The mean of this number is 0.05. Hence, this back-of-the-envelope calculation suggests that, in this setting, only about 10% of all possible risk-sharing relationships which could have involved a family or friendship tie, actually did involve such a tie. Hence, while one might expect kinship and friendship to be major influences in partner choice, there is strong evidence that risk attitude is the more prominent consideration when the partner is being chosen specifically for the purposes of informal insurance. In particular, the family and friendship tie is likely to influence the pool of potential partners (because individuals are less likely to know the risk attitudes of strangers, or to trust them), but the choice of partner from this pool for the purposes of insurance is primarily determined by risk attitudes.

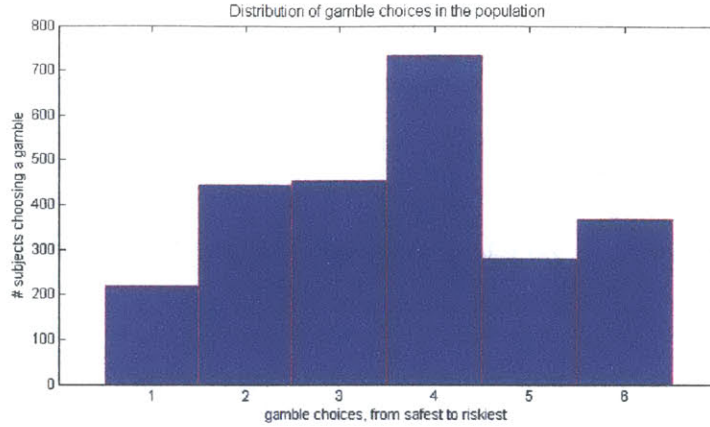
Heterogeneity in Risk Aversion

The second key piece of the model is heterogeneity in risk attitudes across individuals. There is plenty of evidence that people are risk-averse and that they are heterogeneous in their risk-aversion. Experiments which elicit risk attitudes by asking subjects to choose from a set of gambles differing in riskiness find much variation in gamble choice. For example, Harrison et al. (2010) asked 531 experimental subjects drawn from India, Ethiopia, and Uganda to choose a gamble from a set of gambles varying in riskiness (a riskier gamble has higher mean but correspondingly higher variance), in a similar spirit as the seminal study by Binswanger (1980), and estimated the density of CRRA risk attitudes:



It's clear that there is a substantial amount of variation, and almost every point in the $[0, 1]$ range is represented.

In another experiment involving over 2,000 people living in 70 Colombian communities, where 66% live in rural areas, Attanasio et al. (2012) observes the following distribution (gamble 1 is the safest gamble, while gamble 6 is the riskiest):



Chiappori et al. (2010) use two distinct methods to measure heterogeneity in risk preferences in Thai villages, where these villages are spread across several regions in Thailand. The first method is based off the co-movement of individual consumption with aggregate consumption, and the second is based off of optimal portfolio choice theory. Using both methods, they find substantial heterogeneity in risk attitudes in each village. Moreover, this heterogeneity varies across villages and regions. The following table reports the means of risk tolerance for each of 16 villages, and the test statistic for heterogeneity:

village	households	risk aversion γ_i			risk tolerance $1/\gamma_i$		
		mean	χ^2	p-value	mean	χ^2	p-value
<i>Chachoengsao</i>							
2	13	2.00	277.29	0.0000	1.56	3543.60	0.0000
4	21	0.79	78.44	0.0000	2.47	1646.42	0.0000
7	6	0.98	6.69	0.3509	1.28	32.21	0.0000
8	14	0.61	31.11	0.0053	5.11	7986.64	0.0000
<i>Buriram</i>							
2	18	0.62	12.54	0.8184	2.97	368.59	0.0000
10	8	0.34	5.87	0.6618	4.02	147.64	0.0000
13	10	0.41	14.27	0.1610	7.61	2255.00	0.0000
14	15	0.84	73.55	0.0000	3.55	4209.49	0.0000
<i>Lop Buri</i>							
1	19	1.20	96.08	0.0000	1.36	1011.17	0.0000
3	8	2.12	348.07	0.0000	1.33	3981.73	0.0000
4	27	1.40	173.59	0.0000	1.29	2061.54	0.0000
6	24	1.82	485.27	0.0000	1.29	3074.97	0.0000
<i>Sisaket</i>							
1	22	0.43	21.94	0.4633	3.78	457.10	0.0000
6	34	0.78	117.07	0.0000	1.85	2010.67	0.0000
9	22	0.76	33.96	0.0495	3.24	2141.48	0.0000
10	13	0.47	9.68	0.7199	2.90	36.03	0.0006
<i>pooled</i>							
-	274	0.98	1358.43	0.0000	2.64	77568.89	0.0000

Again, it is clear that there is widespread variation in the degree of risk aversion across households.

Heterogeneity in Risky Activities and Settings

Finally, agents in a given setting have a wide variety of investment options and household decisions to make, which vary in riskiness. For example, a farmer must choose a spatial distribution of his plots, what lumpy purchases to make (e.g. bullocks), and when and how to plant his crop. A microentrepreneur must decide what kind of business he wants to start. Parents must decide how to invest household resources, and whom their children will marry. Individuals face a diversity of choices, and how much diversity, as well as the relative riskiness of one decision compared to another, varies across settings.

For example, Rosenzweig and Binswanger (1993) consider the equilibrium crop portfolio choices of heterogeneously risk-averse farmers living in six ICRISAT villages located across three distinct agroclimactic regions in India. The first region is characterized by low levels of erratically distributed rainfall and soils with limited water storage capacity (this is the riskiest environment), the second region by similarly erratic rainfall and irrigation but better soil storage capacity, and the third region by low levels of more reliable rainfall with reasonable soil storage capacity (this is the safest environment). The principal crops grown are sorghum, pigeon peas, pearl millet, chickpeas, and groundnuts, and their yield distributions vary across environment. They show that differences in risk aversion do translate into differences in choice of risky investments. Individuals are influenced by risk-reduction when choosing income streams, particularly in response to limitations on *ex post* consumption-smoothing, and the degree to which they are influenced depends on their risk aversion.

Dercon (1996) also studies the variation in riskiness of agricultural investment decisions by heterogeneously risk-averse rural households. His data is drawn from Tanzania, a country with very underdeveloped credit markets (in 1989, only 5% of commercial bank lending went to the private sector, and less than 10% of this lending went to individual farmers). The UN Food and Agriculture Organization provides an interesting look at the vast heterogeneity in crop yield distributions and equilibrium crop mix across regions in Tanzania in 1998. The following table shows the area, yield, and production of each of five crops across ten agroclimactically heterogeneous regions in Tanzania³⁰:

³⁰Of course, in addition to levels and fluctuations of crop yields, farmers care about the levels and fluctuations of

Table 2: Tanzania - 1998 Vuli Crop Production in Cereal Equivalent by Region

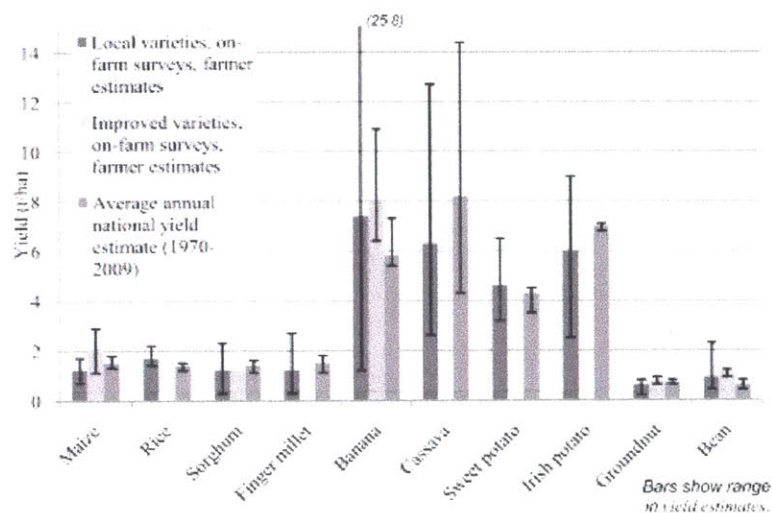
Region	Maize			Sorghum			Paddy			Millet			Wheat		
	Area ('000 ha)	Yield (kg/ha)	Production (tonnes)	Area ('000 ha)	Yield (kg/ha)	Production (tonnes)	Area ('000 ha)	Yield (kg/ha)	Production (tonnes)	Area ('000 ha)	Yield (kg/ha)	Production (tonnes)	Area ('000 ha)	Yield (kg/ha)	Production (tonnes)
Mara	13.0	1 700	22.1	20.0	1 300	26.0	1.0	1 600	1.6	0.5	1 100	0.6	0.0	0	0.0
Arusha	50.0	1 600	80.0	10.0	1 200	12.0	2.6	2 300	5.98	3.4	1 000	3.4	14.0	1 700	23.8
Kilimanjaro	30.0	1 500	45.0	1.5	1 300	2.0	3.2	2 500	8	-	-	-	1.2	1 400	1.7
Tanga	20.0	1 500	30.0	-	-	-	2.1	1 900	3.99	-	-	-	0.7	1 000	0.7
Morogoro	40.0	1 400	56.0	8.0	1 200	9.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Mbeya	20.0	1 600	32.0	0.0	0.0	0.0	1.0	1 800	1.8	0.0	0.0	0.0	0.0	0.0	0.0
Coast/DSM	11.0	1 100	12.1	-	-	-	-	-	-	-	-	-	-	-	-
Kagera	51.0	1 500	76.5	2.7	1 300	3.5	0.1	2 000	0.2	-	-	-	-	-	-
Kigoma	40.0	1 800	72.0	3.0	1 500	4.5	3.0	2 000	6	0.1	1 000	0.1	-	-	-
Mwanza	100.0	1 200	120.0	110.0	1 100	121.0	10.0	1 800	18	1.0	1 000	1.0			
Total	375		545.7	155.2		178.6	23.0		45.6	5.0		5.1	15.9		26.2

Unfortunately, this table excludes estimates of the *variance* of yield of each of these crops across regions. Dercon (1996) provides a discussion of this in his paper. He describes a multiplicity of soils and irrigation systems in Tanzania, which support different crops. Paddy, a crop which can yield a high return, is restricted only to specific soils and areas close to a river, and is the least drought and locust resistant. Despite the potential for high returns, only 11% of the total cultivation sample grew paddy. On the other hand, sorghum yields only a low-moderate return, but all soils can sustain it, and it is more resistant to drought and pests. Even though it had a lower mean return, it was grown by all but two households in the sample.

Uganda and Ethiopia are similar to Tanzania in the set of crops grown, though the actual crop mix grown differs due to differences in environmental conditions. An IFPRI report from 2011 estimating crop yields in Uganda provides a useful illustration of how the variance of crop yields differs across crops, and the typical relationship of the variance with the mean:

crop prices, as they care ultimately about the distribution of profits.

Figure A.1—Average yield and yield range of selected unfertilized crops observed in on-farm surveys (farmer recall) in Uganda, by variety, type, and average of annual national yield estimates (1970–2000)



Source: Compiled by authors based on database of crop yields in Uganda.

There is a clear positive relationship between mean yield and variance of yield. Groundnuts have low mean yields and correspondingly low fluctuation of yields, making it a "safer" crop, while banana has much higher mean yields but correspondingly higher fluctuation of yields, making it a "riskier" crop.

Abebe et al. (2010) provide a similar graphic for Ethiopia:

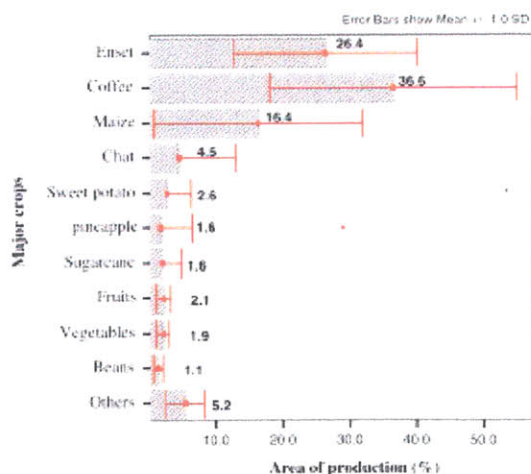


Fig. 3 Mean area share (in percentage of the farm area) of the major crops in farms. All 144 farms are used here. *Error bars* indicate one standard deviation

(Enset is a type of banana.)

Again, we get a general sense that higher mean yield crops have a higher variance of yield, while lower mean yield crops have a smaller variance of yield. Comparing across Uganda and Ethiopia, we see that maize is a safer crop relative to sweet potato in Uganda, but the opposite is true in Ethiopia. Thus, the same set of crops have very different yield distributions in different settings, and furthermore, each crop's relative riskiness with the other crops also varies across setting.

Chapter 2

Risk, Incentives, and Contracting Relationships

2.1 Introduction

Risk is an unavoidable feature of life, but its pervasiveness and the methods which exist to manage it differ in developed and in developing countries. People in developing countries face high levels of risk: disease is widespread, the climate is punishing, and occupations are hazardous (Dercon (2005), Fafchamps (2008)). Instead of steady salaries, incomes are typically highly variable and depend on factors beyond individual control. For example, a farmer's livelihood is bound to the whims of nature, the vagaries of health, and the caprice of crop prices.

In addition, people in developing countries are more vulnerable to the risks they face. They live close to or at subsistence levels, and thus have no buffer with which to cushion negative shocks. Moreover, the formal insurance and credit institutions, credibly backed and enforced by stable governance and legal systems, which are available to people in developed countries, are notably absent in developing environments (Dercon (2005), Fafchamps (2008)).

Consequently, the poor must depend on creative ways of incorporating risk protection into their interactions with each other (Alderman and Paxson (1992), Morduch (1995), Dercon (2005), Fafchamps (2008)). For example, two farmers working together in a cropping group to grow a risky harvest could smooth each other's consumption by agreeing to a sharing rule of the income from the realized harvest. These farmers could also try to establish sharing rules with outsiders, but committing to a division of pooled income becomes prohibitively difficult if the contracting parties

do not jointly observe or cannot easily verify income. As a result, the two farmers have compelling reasons to adapt their working partnership with each other to accommodate risk concerns. This "costly state verification" (Townsend (1979)) often causes subsets of people ostensibly matched for other purposes to incorporate risk management into their existing relationships. (See Townsend and Mueller (1998) for examples of costly state verification and informal insurance in the Indian village of Aurepalle.) Hence, the absence of formal institutions not only induces individuals to rely on their relationships with each other, it also causes these relationships to address *layers* of needs. These relationships and the arrangements within them operate outside of formal insurance and credit channels and can therefore be thought of as "informal institutions". Indeed, Bardhan relates in his 1980 paper on interlocking agrarian factor markets: "Generalizing from his experience with the hill peasants of Orissa, Bailey (1966) notes: 'The watershed between traditional and modern society is exactly this distinction between single-interest and multiplex relationships.'"

In this paper, I study the implications of this multidimensionality of interpersonal relationships arising from the absence of formal institutions by focusing specifically on the influence of missing formal insurance institutions on the equilibrium formation of productive partnerships between risk-averse people and subject to one-sided moral hazard. A classic example is sharecropping: a farmer lives on and crops a plot belonging to a landowner and pays rent as a share of the realized harvest. Sharecropping has thrived in different parts of the world for centuries, and continues to be prevalent in parts of the world today, for example in rice farms in Bangladesh (Akanda et. al. (2008)) and in Madagascar (Bellemare (2009)). If inputs (including effort) are difficult to monitor and are noncontractible, however, the contract that provides first-best incentives for the farmer should be the contract that makes the farmer the residual claimant (the farmer should pay a fixed rent). So why the share contract? Stiglitz (1974) suggested that the share contract emerged because employment relationships also provided protection from risk, in the absence of formal insurance: while a fixed rent contract would induce the right incentives, it would force the agent to bear the full risk of the stochastic harvest. A fixed wage contract would fully insure the agent, but the agent would have no incentive to exert costly effort.

However, while the effect of the tradeoff between incentive and insurance provision on a contract between a given pair of principals and agents has been studied extensively, the effect of this tradeoff on the *formation* of employment relationships has received far less attention. This latter analysis is essential for rigorously understanding the true strength of informal insurance—that is, the level of risk-sharing achieved within the network of equilibrium employment relationships which emerges

in the absence of formal insurance and credit institutions, and not merely the risk-sharing achieved with a single, isolated group of individuals.

To perform this analysis rigorously, I build a model of endogenous one-to-one matching between heterogeneously risk-averse principals and agents who face a classic moral hazard problem. Principals each own a unit of physical capital, but have infinite marginal cost of effort, while agents own zero units of physical capital, but have the same finite marginal cost of effort. Physical capital must be combined with human capital in order to produce output. For example, in a sharecropping setting, landowners would be principals, farmers would be agents, and land and farming expertise would need to be combined to produce any sort of harvest.

The distribution of output depends additively on the unobservable and noncontractible effort exerted by the agent, as well as on the level of riskiness of the environment. More specifically, an increase in effort exerted by the agent leads to an increase in mean output but has no impact on the variance, while the level of riskiness of the environment determines the variance of output. A matched principal and agent can commit *ex ante* to a return-contingent sharing rule of their jointly-produced output.

The difficulty of characterizing the equilibrium wage schedule in moral hazard models in which both the principal and the agent are risk-averse (and have differing risk attitudes) is well-known. Instead of using a Holmstrom and Milgrom (1987) type story where the principal observes some coarser aggregate of output than the agent does to justify linear contracts in a CARA-Normal framework, I develop a model of one-sided moral hazard in which principals and agents have CARA utility and returns are distributed Laplace. The Laplace distribution has two key features: first, it resembles the normal distribution (for example, it is symmetric about the mean), but has fatter tails, and second, its likelihood ratio is a piecewise constant with discontinuity at the mean. (Please see Appendix 1 for further details of the Laplace distribution.) This framework is optimally suited to analyzing the formation of equilibrium networks of relationships subject to one-sided moral hazard and risk-sharing, since the equilibrium wage cleanly separates incentive and insurance provision. The equilibrium wage schedule is piecewise linear with discontinuity at the mean, where the linearity at output levels away from the mean captures efficient risk-sharing between the risk-averse principal and agent, and the discontinuity at the mean captures incentive provision. (Please see Appendix 1 for a detailed proof and discussion of this result.)

I characterize sufficient conditions for the unique equilibrium matching between principals and agents to be negative-assortative in risk attitude (the i^{th} least risk-averse principal works with the

i^{th} most risk-averse agent, and so on), and for the unique equilibrium matching to be positive-assortative in risk attitude (the i^{th} least risk-averse principal works with the i^{th} least risk-averse agent). Intuitively, the equilibrium risk composition of partnerships is determined by the tradeoff between incentive and insurance provision in the following way: a less risk-averse principal who hires a more risk-averse agent can charge that agent a risk premium for insurance, but the more that the principal insures the agent, the less effort that agent will exert. Moreover, a more risk-averse agent has a higher effective marginal cost of effort in the first place (because a more risk-averse agent is "more concerned" about the possibility of exerting high effort but being unlucky and realizing a low output draw).

I show that the key determinants of equilibrium group composition are: the curvature of the cost of effort function, the riskiness of the environment, and the distribution of risk types in the economy. The curvature of the cost of effort function influences the equilibrium match because it characterizes the tradeoff between insurance and incentive provision across partnerships of different risk attitudes. Consideration of an extreme example provides the intuition. Suppose the cost of effort function were "infinitely convex". Then, regardless of risk attitude, any agent would exert near zero effort. But this means that no agent is cheaper to incentivize than any other agent—that is, all agents are equally expensive to incentivize. Hence, from the principals' perspective, incentive provision is effectively held fixed across agents of different risk attitudes, and the matching becomes driven purely by insurance provision. So, the equilibrium match will be negative-assortative in risk attitudes, since the less risk-averse principals are differentially more willing to provide insurance to the more risk-averse agents, who are differentially more willing to purchase it. As the cost of effort function becomes less extreme, however, the less risk-averse agents become notably cheaper to incentivize than the more risk-averse agents, and positive-assortative matching may arise.

The riskiness of the environment is a particularly interesting determinant of the equilibrium match. It influences the matching in two important ways: first, the safer the environment is, the lower the premium for insurance. Second, the safer the environment, the more informative output is as a signal of agent effort. In other words, insurance provision becomes less differentiated across agents of different risk attitudes when the environment is safer, because even the more risk-averse agents don't need so much insurance, while incentive provision more sharply differentiates across agents of different risk attitudes when the environment is safer, because there is less room for hidden action.

Finally, the distributions of risk types in the group of principals and in the group of agents

affect the equilibrium match¹. This is because effort is supplied only by the agent, and not by the principal—hence, the risk attitude of a principal has a different effect on equilibrium effort than does the risk attitude of an agent. The one-sidedness of moral hazard generates an asymmetry in the model which makes the distribution of risk types important for matching: equilibrium partnerships depend both on "within group heterogeneity", that is, how much risk attitudes vary among principals, and how much risk attitudes vary among agents, as well as "across group heterogeneity", that is, how different the risk attitudes of principals are from the risk attitudes of agents.

Loosely, the features of the environment in which negative-assortative matching emerges as the unique equilibrium are: a cost of effort function which is either close to linearity or is extremely convex, a highly risky environment (output is a very noisy signal of effort, and the premium paid for insurance is high), and a group of principals who are distinctly more risk-averse than the agents. By contrast, the features of the environment where positive-assortative matching emerges as the unique equilibrium are: a "moderately convex" cost of effort function, a safe environment (output is a quite precise signal of effort, and the premium paid for insurance is low), and a group of principals who are distinctly less risk-averse than the agents. To be slightly more specific, the distribution of risk types influences the equilibrium match by affecting the bounds which delineate these "close to linear", "moderatey convex", and "extremely convex" descriptors of the curvature of the cost of effort function.

After establishing the main theoretical results, I work through a numeric example which demonstrates the policymaking importance of understanding the risk composition of the equilibrium network of partnerships: evaluating policy requires understanding how people will re-optimize upon its introduction, and understanding how people will re-optimize in response to the introduction of a policy requires understanding how people optimize in the status quo. That is, in order to predict what effect the introduction of a formal institution would have, it is necessary first to have a complete picture of the operations of the equilibrium informal institutions. For example, I discuss the impact of introducing formal insurance. I show that such a policy leads to a "crowding out" of informal insurance, and in particular may leave the providers of informal insurance worse off. This framework of endogenous matching is exactly suited to analyzing concretely and rigorously the "crowding out" effect, which is often discussed in the literature.

¹This is in contrast with the distribution-free result from Wang (2012a) which studied the impact of the trade-off between two informal insurance techniques (income-smoothing and consumption-smoothing) on the equilibrium formation of relationships.

A small body of existing literature examines the relationship between the insurance and incentive provision tradeoff and the endogenous formation of contracting relationships under one-sided moral hazard. Legros and Newman (2007) studied more generally the problem of endogenous matching under nontransferable utility. They present techniques to characterize stable matchings in nontransferable utility settings by generalizing the Shapley and Shubik (1972) and Becker (1974) supermodularity and submodularity conditions for matching under transferable utility. Under non-transferable utility, the indirect utility of each member of the first group given a partnership with each member of the second group can be calculated, fixing the second member's level of expected utility at some level v . Then, this indirect utility expression, which depends on both members' types and v , is analyzed for supermodularity and submodularity in risk types.

Franco et. al. (2011) consider a risk-neutral principal who requires two (risk-neutral) agents to operate her machinery, where agents differ in their marginal cost of effort (high or low). The principal must decide what teams of agents to form—low marginal cost with low marginal cost (positive-assortative), or low with high (negative-assortative). Moral hazard is double-sided—the key new force behind the matching decision is the principal's inability to condition compensation on an individual agent's contribution. The main result is that the super/submodularity of the production technology in workers' inputs no longer drives assortative matching in the presence of this double-sided moral hazard. When the production technology is modular, so that in the absence of moral hazard there is no matching prediction, Franco et. al. show that the presence of moral hazard can still lead to positive- or negative-assortative matching, depending on the optimal compensation scheme, which depends on the types and output of the team. For example, if types exerting higher input are rewarded more according to the scheme, then negative matching is optimal, to reduce the likelihood of "accidental payment", that is, paying an individual who exerts low effort for the high effort exerted by his partner. When the technology exhibits complementarities, a scenario is described where increasing complementarity does not lead to positive-assortative matching because of the double-sided moral hazard.

Serfes (2008) studies a setting where risk-neutral principals owning exogenously-assigned projects match with risk-averse agents, where the projects of principals vary in riskiness (riskier projects have higher mean and higher variance), and the risk aversion of agents also varies. A principal and an agent jointly produce output, where output depends additively on unobservable and noncontractible effort exerted by the agent. Serfes shows that the equilibrium match is often not globally assortative in risk attitude, and that the relationship between the riskiness of the environment and

the power of the contract can be ambiguous, due to the endogeneity of matching.

Importantly, the approach to studying the tradeoff between incentive and insurance provision in this paper is markedly different from approaches in papers such as the one by Serfes, where principals are assumed to be risk-neutral, but own projects which vary in riskiness. While such a model generates a tractable equilibrium wage (because risk-neutrality ensures linearity), the endogenous assignment problem of heterogeneously risk-averse agents to risk-neutral principals with heterogeneously risky projects is really answering the question, "What risky project is a risk-averse agent assigned to when the output of the project depends on the noncontractible effort of the agent, and a risk-neutral insurer is available who sells insurance?" That is, this approach can be thought of as focusing on the impact of a formal institution on equilibrium activities undertaken in a village, when there is some sort of monitoring problem. By contrast, my paper builds a model of endogenous matching between heterogeneously risk-averse principals and heterogeneously risk-averse agents, where only one project is available (for example, wheat is the only crop that can be grown). This approach focuses on the emergence, structure, and performance of informal insurance institutions in the status quo. Hence, the first question that this paper answers is, "How well-insured are heterogeneously risk-averse individuals when a lack of formal institutions pushes their interpersonal relationships to address multiple needs, including the need for risk protection?"

Chiappori and Reny (2006) and Schulhofer-Wohl (2006) both study a model of endogenous matching between heterogeneously risk-averse individuals under pure consumption-smoothing, without moral hazard. Any matched pair faces the same exogenous risk, but a matched pair is able to commit to a return-contingent sharing rule given that risk, where output does not depend on any actions of the individuals.

Schulhofer-Wohl and Chiappori and Reny find that negative-assortative matching arises as the unique equilibrium. Instead of using the approach of Legros and Newman, which is elegant but may be intractably algebraic, Schulhofer-Wohl identifies a transferable utility representation of his model, and applies the standard Shapley and Shubik supermodularity conditions. (I take this approach as well.) The key insight is that a less risk-averse man is differentially happier than a more risk-averse man to provide insurance for the most risk-averse woman, who is differentially happier than a less risk-averse woman to pay for it.

A number of papers have attempted to empirically detect the inverse relationship between the riskiness of the environment and the power of the contract predicted by the basic principal-agent model with risk and moral hazard. For example, Allen and Lueck (1992) study sharecropping

relationships in the American Midwest in 1986, and observe that the strength of dependence of a sharecropper's rent contract on realized harvest is *positively* correlated with the riskiness of the crop grown. From this, they conclude that risk is not a problem for farmers, since risk considerations do not appear to influence contract design. However, if more risk-averse farmers work for less risk-averse landowners, cultivating safer crops under low-powered contracts, while less risk-averse farmers work for more risk-averse landowners, cultivating riskier crops under high-powered contracts, the same empirical observations would emerge, but risk concerns would be playing a significant role in contract design, through the unaccounted-for channel of contracting partner choice. Alternatively, we expect formal institutions to be stronger in the United States—hence landowner-farmer relationships may not be as multidimensional as they are in developing countries. This would imply that this is not the right dataset to test for the theoretically-predicted relationship between risk and incentives.

The remainder of the paper proceeds as follows. The next two sections present the model and the results. Section 4 works through a hypothetical policy example and shows that the evaluation of policy may change drastically if it accounts for the endogenous network response. Section 5 concludes.

2.2 The Model

In this section, I introduce a framework designed to analyze the formation of equilibrium contracting relationships under one-sided moral hazard when the contract balances insurance provision with incentive provision.

The framework consists of the following elements:

The population of agents: the economy is populated by two groups of agents, $G1$ and $G2$, where $|G1| = |G2| = Z$, Z a finite, positive integer. Call members of $G1$ "principals" and members of $G2$, "agents". All principals own one unit of physical capital, but have infinite marginal cost of effort; all agents lack physical capital but have finite and identical marginal costs of effort. The cost of effort for all agents is $c(a)$, $c(a) > 0$, $c'(a) > 0$, $c''(a) > 0$.

Principals and agents both have CARA utility, $u(x; r) = -e^{-rx}$, where individuals differ in their degree of risk aversion. Let r_1 represent a principal, and r_2 represent an agent. Principals and agents are identical in all other aspects. There are no assumptions on distributions of risk types.²

²Of course, in reality, types are multidimensional, and matching decisions are not exclusively based on risk atti-

The risky environment: A principal-agent partnership can only produce positive output if one unit of physical capital is combined with human capital. For example, a landowner owns land, and a farmer has agricultural experience and skill, and a successful harvest requires the landowner and farmer to combine their capital. Output is given by $R|a = \gamma a + \varepsilon$, where the riskiness of the environment is captured by $\varepsilon \sim f_\varepsilon$, an exogenously-given, well-defined, differentiable probability distribution function with support on $(-\infty, \infty)$. The effort of an agent, a , increases the mean but doesn't affect the variance of returns.

Information and commitment: All agents know each other's risk types. In the first-best environment, an agent's effort is both observable and contractible. In the second-best environment, the agent's effort is neither observable nor contractible.

A given matched pair (r_1, r_2) observes the realized output of their partnership, and is able to commit *ex ante* to a return-contingent sharing rule $s(R_{p_{12}})$, where $R_{p_{12}}$ is the realized return of (r_1, r_2) 's joint project p_{12} . More precisely, $s(R_{p_{12}})$ specifies the wage paid to the agent r_2 when the realized return is $R_{p_{12}}$, where $s : \mathbb{R} \rightarrow \mathbb{R}$ (there are no limited liability assumptions). In order to be *feasible*, the income the principal r_1 receives must be less than or equal to $R_{12} - s(R_{p_{12}})$. Since all individuals have monotonically increasing utility, r_1 's share will be equal to $R_{12} - s(R_{p_{12}})$.

The equilibrium: An equilibrium is:

1. A match function $\mu(r_1) = r_2$, where $\mu(\cdot)$ assigns each r_1 to at most one agent r_2 , and distinct people have distinct partners.

Moreover, the matching pattern described by $\mu(\cdot)$ must be stable. That is, it must satisfy two properties:

- (a) *No blocks*: no unmatched principal and agent should be able to write a feasible wage contract such that both of them are happier with each other than they are with the partners assigned to them by μ .

tudes. It is worth noting that the model can account for this. For example, kinship and friendship ties are important, in large part because of information (they know each other's risk types), and commitment (they trust each other, or can discipline each other). Kinship and friendship ties would enter into this theory in the following way: an individual would first identify a pool of feasible risk-sharing partners. This pool would be determined by kinship and friendship ties, because of good information and commitment. Following this, individuals would choose risk-sharing partners from these pools. This choice would be driven by risk attitudes, as addressed in this benchmark with full information and commitment.

Thus, this theory can be thought of as addressing the stage of matching that occurs after pools of feasible partners have been identified.

(b) *Individual rationality*: each agent must receive a higher expected utility from being in the match $\mu(\cdot)$ than from remaining unmatched.

2. A set of sharing rules and effort choices by the agents, one sharing rule and one effort choice for each matched pair. The agent chooses an effort level which is optimal for her—she should not be able to choose a different effort level and become better off. Furthermore, no pair should be able to choose a different sharing rule which leaves both partners weakly better off, and at least one partner strictly better off (the agent chooses effort optimally in response to the sharing rule).

Matching patterns: It will be helpful to introduce some matching terminology. Suppose the people in $G1$ and in $G2$ are ordered from least to most risk-averse: $\{r_1^j, r_2^j, \dots, r_Z^j\}$, $j \in \{1, 2\}$. Then "positive-assortative matching" (PAM) refers to the case where the i^{th} least risk-averse person in $G1$ is matched with the i^{th} least risk-averse person in $G2$: $\mu(r_i^1) = r_i^2$, $i \in \{1, \dots, Z\}$. On the other hand, "negative-assortative matching" (NAM) refers to the case where the i^{th} least risk-averse person in $G1$ is matched with the i^{th} most risk-averse person in $G2$: $\mu(r_i^1) = r_{Z-i+1}^2$, $i \in \{1, \dots, Z\}$. To say that the unique equilibrium matching pattern is PAM, for example, is to mean that the only μ which can be stable under optimal within-pair sharing rules and projects is the match function which assigns agents to each other positive-assortatively in risk attitudes.

2.3 Results

2.3.1 The First-Best

It will be useful to begin by solving the first-best problem, when the agent's effort is observable and contractible.

The first step to characterizing the equilibrium network of relationships is to characterize what happens in a given relationship. Suppose principal r_1 is matched with agent r_2 , and that the returns of the risky project R are distributed according to some general density function f :

$$R|a = \gamma a + \varepsilon, \varepsilon \sim f(\varepsilon)$$

Assume the random variable R has a well-defined cumulant generating function³.

³Recall that the cumulant-generating function is the log of the moment-generating function.

Denote the agent's share of realized output R by $s(R)$. Then, the pair's equilibrium sharing rule, given that r_2 receives expected utility at least $-e^{-v}$ for some fixed level v , solves the following problem:

$$\begin{aligned} \max_{a, s(R)} \int_{-\infty}^{\infty} -e^{-r_1[R-s(R)]} f(R-\gamma a) dR \text{ s.t.} \\ (IR) \int_{-\infty}^{\infty} -e^{-r_2[s(R)-c(a)]} f(R-\gamma a) dR \geq -e^{-v} \end{aligned}$$

The equilibrium sharing rule and effort are described in the following lemma.

Lemma 12 *The optimal first-best contract of a principal-agent pair (r_1, r_2) , where $s(R)$ denotes the agent's share, is:*

$$\begin{aligned} s_{FB}^*(R) &= \frac{r_1}{r_1 + r_2} R + K_{FB}^*(r_1, r_2, v) \\ K_{FB}^*(r_1, r_2, v) &= \frac{1}{r_2} v + c(c'^{-1}(\gamma)) + \frac{1}{r_2} \log \left(\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R} f(R - \gamma c'^{-1}(\gamma)) dR \right) \\ a_{FB}^* &= c'^{-1}(\gamma) \end{aligned}$$

(The proof is in Appendix 2.)

The first-best contract has several notable features. First, equilibrium effort level is the same in any possible pair—this is because effort is contractible. Hence, any principal-agent pair chooses the effort which "maximizes the pie", and then efficiently shares the risk of that pie. The effort which "maximizes the pie" is the effort level which equates the marginal benefit of effort (the marginal impact on mean output, γ) with the marginal cost of effort exertion, $c'(\cdot)$.

Second, the equilibrium wage is linear. This is unsurprising—again, effort is contractible, so the sharing rule needs only to provide insurance, and not incentives. Moreover, the less risk-averse individual receives a share that is more heavily dependent on output realization.

Now that we have characterized the optimal sharing rule and equilibrium effort within a matched pair r_1 and r_2 , we can solve for the equilibrium network of relationships. Intuitively, since we know that any possible matched pair chooses the same effort level, we expect negative-assortative matching to arise as the unique equilibrium. This is because we know from Schulhofer-Wohl and Chiappori and Reny that endogenous matching under pure *ex post* risk management results in unique negative-assortative matching (there is no moral hazard, and no scope for *ex ante* risk management—matched

pairs are not able to choose what risk they face). Negative-assortative matching arises because the least risk-averse individuals are differentially willing to provide insurance, while the most risk-averse individuals are differentially willing to pay for it.

To formalize this intuition, we need to identify a method for characterizing the equilibrium match. A challenge is posed by the heterogeneity of risk-aversion in agents, which makes this a model of matching under nontransferable utility. That is, the amount of utility experienced by an agent with risk aversion r_1 from consuming one unit of output differs from the amount of utility an agent with risk aversion r_2 experiences from one unit of output. Thus, we cannot directly apply the Shapley and Shubik (1962) result on sufficient conditions for assortative matching in transferable utility games.

It will be helpful to review briefly that environment and result. Consider a population consisting of two groups of risk-neutral workers, where all workers have utility $u(c) = c$. Let a_1 denote the ability of workers in one group, and a_2 denote the ability of workers in the other group. The production function is given by $f(a_1, a_2)$, which can be thought of as: "the size of the pie generated by matched workers a_1 and a_2 ". Then, $\frac{d^2 f}{da_1 da_2} > 0$ is a sufficient condition for unique positive-assortative matching, and $\frac{d^2 f}{da_1 da_2} < 0$ is a sufficient condition for unique negative-assortative matching.

My approach here will be to identify the function in this model of nontransferable utility which is analogous to the Shapley and Shubik production function $f(a_1, a_2)$. In Proposition 2 below, I prove that *expected utility* is transferable in this model—instead of thinking about moving "*ex post*" units of output between agents, we should instead think about moving "*ex ante*" units of expected utility. I show that the sum of the certainty-equivalents $CE(r_1, r_2)$ of a given matched pair (r_1, r_2) is the analogy to the joint output production function in the transferable utility problem. The sum of the certainty-equivalents of a matched pair is "the size of the expected utility pie generated by matched agents r_1 and r_2 ", and sufficient conditions for positive-assortative and negative-assortative matching correspond to conditions for the supermodularity and submodularity of $CE(r_1, r_2)$ in r_1, r_2 .

More technically, expected utility is transferable in this model because the expected utility Pareto possibility frontier for a pair (r_1, r_2) is a line with slope -1 under some monotonic transformation.

Proposition 13 *Expected utility is transferable in this model.*

Proof. Using the optimal sharing rule and equilibrium effort from Lemma 1, we can write the expressions for the certainty-equivalent of a principal r_1 and of an agent r_2 who are matched with each other:

$$CE_{r_1} = -\frac{v}{r_2} - \left(\frac{1}{r_1} + \frac{1}{r_2} \right) \log E \left[e^{-\frac{r_1 r_2}{r_1 + r_2} R} | a^* = c'^{-1}(\gamma) \right] - c(c'^{-1}(\gamma))$$

$$CE_{r_2} = \frac{v}{r_2}$$

Hence, it is clear the cost to the principal r_1 of increasing the certainty-equivalent of her agent r_2 by one unit is exactly one unit (and vice versa). That is, expected utility is transferable in the model. ■

This tells us that the sum of certainty-equivalents for a matched principal and agent in this model can be thought of as the function which is analogous to the joint output function in the Shapley and Shubik transferable utility setting:

$$CE(r_1, r_2) = -\left(\frac{1}{r_1} + \frac{1}{r_2} \right) \log E \left[e^{-\frac{r_1 r_2}{r_1 + r_2} R} | a^* = c'^{-1}(\gamma) \right] - c(c'^{-1}(\gamma))$$

It will be helpful to make two observations at this point. First, we can define the *representative risk aversion* of a matched pair (r_1, r_2) :

$$\hat{H}(r_1, r_2) = \frac{r_1 r_2}{r_1 + r_2}$$

Importantly, we can see that the sum of certainty-equivalents of a matched pair depends *only* on representative risk aversion. In other words, a matched pair (r_1, r_2) acts as a single individual with CARA utility and absolute risk aversion $\hat{H}(r_1, r_2)$.

And correspondingly, we can define the reciprocal of representative risk aversion, the *representative risk tolerance*:

$$\begin{aligned} \tilde{H}(r_1, r_2) &= \frac{1}{\hat{H}(r_1, r_2)} \\ &= \frac{1}{r_1} + \frac{1}{r_2} \end{aligned}$$

Thus, the sum of certainty-equivalents written as a function of representative risk tolerance is:

$$CE(\tilde{H}) = -\tilde{H} \log E \left[e^{-\frac{1}{\tilde{H}} R} | a^* = c'^{-1}(\gamma) \right] - c(c'^{-1}(\gamma))$$

Intuitively, the joint expected utility pie of a matched pair is some transformation of the representative individual's expected utility from facing the stream of returns R , minus the cost of effort, where effort in the first-best is independent of any risk types.

Furthermore, the transformation of the representative individual's expected utility is a special transformation:

$$K_{R|a^*} \left(t = -\frac{1}{\tilde{H}} \right) = \log E \left[e^{-\frac{1}{\tilde{H}} R} | a^* = c'^{-1}(\gamma) \right]$$

where $K_{R|a^*} \left(t = -\frac{1}{\tilde{H}} \right)$ is the cumulant-generating function (log of the moment-generating function) of the random variable $R|a^*$ evaluated at the negative of the representative risk tolerance.

How does this contribute to our understanding of equilibrium matching in the economy? We know from the Shapley and Shubik assortative matching conditions that a sufficient condition for unique PAM in this setting is supermodularity of $CE(r_1, r_2)$ in r_1, r_2 , and a sufficient condition for unique NAM in this setting is submodularity of $CE(r_1, r_2)$ in r_1, r_2 .

Moreover, we know that:

$$\begin{aligned} \frac{d^2 CE(r_1, r_2)}{dr_1 dr_2} &= \frac{dCE}{d\tilde{H}} \frac{d^2 \tilde{H}}{dr_1 dr_2} + \frac{d^2 CE}{d\tilde{H}^2} \frac{d\tilde{H}}{dr_1} \frac{d\tilde{H}}{dr_2} \\ &= \left(\frac{1}{r_1 r_2} \right)^2 \frac{d^2 CE}{d\tilde{H}^2} \end{aligned}$$

Hence, a sufficient condition for unique PAM is convexity of $CE(\tilde{H})$ in \tilde{H} , while concavity of $CE(\tilde{H})$ in \tilde{H} ensures unique NAM.

Checking the second derivative of $CE(\tilde{H})$ in \tilde{H} yields the first-best matching result.

Proposition 14 *In the first-best model, when effort is observable and contractible, the unique equilibrium matching pattern in risk attitude of principals and agents is negative-assortative.*

Proof. The second derivative of $CE(\tilde{H})$ in \tilde{H} is straightforward to find.

$$CE(\tilde{H}) = -\tilde{H} K_{R|a^*} \left(t = -\frac{1}{\tilde{H}} \right) - c(c'^{-1}(\gamma))$$

$$\begin{aligned}\frac{dCE(\tilde{H})}{d\tilde{H}} &= -K_{R|a^*}\left(t = -\frac{1}{\tilde{H}}\right) - \frac{1}{\tilde{H}}K'_{R|a^*}\left(t = -\frac{1}{\tilde{H}}\right) \\ \frac{d^2CE(\tilde{H})}{d\tilde{H}^2} &= -\frac{1}{\tilde{H}^3}K''_{R|a^*}\left(t = -\frac{1}{\tilde{H}}\right) < 0\end{aligned}$$

since the cumulant-generating function is convex in t .

Hence, the unique equilibrium match in the first-best is negative-assortative. ■

Now that we understand what happens in the first-best, we can investigate the second-best.

2.3.2 The Second-Best

Now, suppose that effort is not observable and not contractible. In this case, the sharing rule of a given principal-agent pair must provide incentives as well as insurance.

To gain traction on this problem, I impose a specific functional form assumption on the distribution of output.

Recall that output given effort a is described by $R|a = \gamma a + \varepsilon$. In the previous subsection, it was assumed that $\varepsilon \sim f_\varepsilon$, a general density function with support on the real line and well-defined cumulant-generating function.

In this subsection, assume that f_ε is a Laplace distribution with mean 0 and exogenously-given variance $V > 0$, where V captures the riskiness of the environment.⁴ (See Appendix 1 for more details on the Laplace distribution.)

The key features of this distribution for the setting of this paper are symmetry. The distribution resembles the normal distribution, but has fatter tails, and the density function is non-differentiable at the mean. Loosely, the fatter tails allow us to avoid the Mirrlees critique of linear contracts—a realized return in the tail of a Laplace distribution is not infinitely precise about effort exerted.

Then, the equilibrium sharing rule of a principal-agent pair (r_1, r_2) , where the agent r_2 is ensured expected utility at least $-e^{-v}$, solves:

$$\max_{s(R)} \int_{-\infty}^{\gamma a} -e^{-r_1(R-s(R))} \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR + \int_{\gamma a}^{\infty} -e^{-r_1(R-s(R))} \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR$$

such that:

$$(IR) : \int_{-\infty}^{\gamma a} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR +$$

⁴ Assume $V < \frac{1}{\max(r_1)} + \frac{1}{\max(r_2)}$, for the problem to be well-defined.

$$\begin{aligned}
& + \int_{\gamma a}^{\infty} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR \geq -e^{-v} \\
(IC) : a \in \arg \max_{a \in (0, \infty)} & \int_{-\infty}^{\gamma a} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR + \\
& + \int_{\gamma a}^{\infty} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR
\end{aligned}$$

The complete equilibrium analysis of this problem can be found in Appendix 1, but I will provide a sketch of the solution here, to provide an understanding of the equilibrium behavior of a matched partnership.

First, we need to address the constraints. The *IR* constraint clearly binds in equilibrium. The *IC* constraint presents more of a challenge. It would be useful to be able to replace the global *IC* constraint with its first-order condition, but none of the existing sufficient conditions for the validity of the first-order approach apply to this model. Rogerson (1985) shows that sufficient conditions for a standard moral hazard model in which the principal may also be risk-averse are: (a) monotone likelihood ratio, and (b) convexity of the distribution function. Condition (a) holds in my model (the likelihood ratio here is a piecewise constant, $-c$ below the mean and c above the mean), but (b) fails—very few standard distribution functions satisfy CDFC. Jewitt (1988) identifies sufficient conditions that weaken CDFC, but for a model with a risk-neutral principal. Additionally, utility is assumed to be additively separable in consumption and effort, whereas it is multiplicatively separable here. A variety of more recent contributions identify sets of conditions that weaken CDFC slightly, at the cost of strengthening other conditions, but none weakens CDFC enough for the Laplace distribution.

So, the validity of the first-order approach must be proved from first principles⁵. Since effort a is chosen from an open set, the optimum will be interior, if it exists. Hence, the first-order condition of the global *IC* constraint is a necessary, though perhaps not sufficient, condition for the optimum.

This means that the first-order problem (the problem with the global *IC* condition replaced by its first-order condition) is a relaxed problem, so that the actual optimum must be a solution of the first-order problem, if it exists, but a solution of the first-order problem is not necessarily the optimum.

Solving the first-order problem yields the following wage schedule:

⁵ Again, a rigorous proof can be found in Wang (2012b).

$$\begin{aligned}
s(R_p < \gamma\hat{a}) &= \frac{r_1}{r_1 + r_2}R - \frac{r_1}{r_1 + r_2}\gamma\hat{a} + c(\hat{a}) + \frac{1}{r_2}v - \frac{1}{r_2}\log\left[\left(1 - \frac{r_1r_2}{r_1 + r_2}V\right)\left(1 + \frac{r_2c'(\hat{a})V}{\gamma}\right)\right] \\
s(R_p > \gamma\hat{a}) &= \frac{r_1}{r_1 + r_2}R_p - \frac{r_1}{r_1 + r_2}\gamma\hat{a} + c(\hat{a}) + \frac{1}{r_2}v - \frac{1}{r_2}\log\left[\left(1 + \frac{r_1r_2}{r_1 + r_2}V\right)\left(1 - \frac{r_2c'(\hat{a})V}{\gamma}\right)\right]
\end{aligned}$$

where r_1 is the risk attitude of the principal, r_2 is the risk attitude of the agent, and \hat{a} is the level of effort "anticipated" by the principal. In equilibrium, the optimal effort chosen by the agent in response to the wage schedule $s(R_p|\hat{a})$ should be $a^* = \hat{a}$: the principal has no incentive to pay for a higher level of effort than she knows will actually be exerted, and the agent has no incentive to exert more effort than he is compensated for.

What is equilibrium effort given this compensation scheme? It can be shown that $a = \hat{a}$ is a stationary point of agent r_2 's expected utility from exerting effort a given wage schedule $s(R|\hat{a})$, for every possible \hat{a} . However, for $\hat{a} > \hat{a}_t$, where \hat{a}_t is some threshold, there will be a second stationary point at $a < \hat{a}$ —because the wage schedule is discontinuous for $\hat{a} \neq c'^{-1}\left(\frac{r_1}{r_1+r_2}\gamma\right)$, if the principal tries to induce a "too-high" level of effort, the agent will profitably deviate to a *discretely* lower level of effort.

More concisely, for $\hat{a} \leq \hat{a}_t$, the unique maximizing level of effort exerted by the agent is $a^* = \hat{a}$; for $\hat{a} > \hat{a}_t$, the unique maximizing level of effort exerted by the agent is $a < \hat{a}$.

Therefore, the equilibrium \hat{a} set by the principal is $\hat{a}^* = \hat{a}_t$, where this threshold is characterized by:

$$\left(c'(\hat{a}_t) - \frac{r_1}{r_1 + r_2}\gamma\right)\left(c'(\hat{a}_t) + \frac{\gamma}{r_2V}\right) = \frac{1}{r_2}c''(\hat{a}_t) > 0$$

Because the agent's expected utility from exerting effort a given wage $s(R|\hat{a}^*)$ is strictly concave in a , it must be that $s(R|\hat{a}^*)$ is in fact the optimum.

Observe that setting $\hat{a}_t = c'^{-1}\left(\frac{r_1}{r_1+r_2}\gamma\right)$ causes the left-hand side of the equation to be 0, while the right-hand side is positive. Since the left-hand side is strictly increasing in \hat{a}_t (the cost function $c(a)$ is strictly convex), it must be that $\hat{a}_t > c'^{-1}\left(\frac{r_1}{r_1+r_2}\gamma\right)$.

Therefore, the wage schedule within a principal-agent pair (r_1, r_2) is piecewise linear: at the anticipated mean level of output, $\gamma\hat{a}$, there is a jump in the wage—realized output levels greater than the mean $\gamma\hat{a}^*$ are rewarded at a discretely higher level than output levels that are below the mean. At output levels away from the anticipated mean, the wage is linear with slope $\frac{r_1}{r_1+r_2}$. Hence,

the equilibrium wage can be cleanly decomposed into insurance provision and incentive provision. The jump in the wage at the mean provides incentives (since the likelihood ratio is a piecewise constant, in some sense knowing whether output is above or below the mean is differentially more informative about effort exertion), and the linearity away from the mean captures risk-sharing.

Using this characterization of equilibrium sharing rule and effort in a given principal-agent pair (r_1, r_2) , we can solve for conditions for unique assortative matching, under a functional form assumption on cost of effort: $c(a) = \eta a^M$, $M > 1$ for convexity.

We use the same trick as in the first-best: expected utility is transferable in the second-best as well.

Proposition 15 *Expected utility is transferable in the second-best.*

Proof. Using the equilibrium sharing rule of a given pair and the characterization of equilibrium effort, we can write the certainty-equivalent of principal r_1 and agent r_2 when matched:

$$\begin{aligned}
CE_{r_1} &= \gamma \hat{a}_{12} - c(\hat{a}_{12}) - \frac{v}{r_2} \\
&\quad - \frac{1}{r_1} \log \left(\frac{1}{2} \left[\frac{1}{\left[1 + \frac{r_2 c'(\hat{a}_{12}) V}{\gamma}\right]^{\frac{r_1}{r_2}} \left[1 - \frac{r_1 r_2}{r_1 + r_2} V\right]^{1 + \frac{r_1}{r_2}}} + \frac{1}{\left[1 - \frac{r_2 c'(\hat{a}_{12}) V}{\gamma}\right]^{\frac{r_1}{r_2}} \left[1 + \frac{r_1 r_2}{r_1 + r_2} V\right]^{1 + \frac{r_1}{r_2}}} \right] \right) \\
CE_{r_2} &= \frac{v}{r_2}
\end{aligned}$$

Hence, it is clear that there is a one-to-one tradeoff in the certainty-equivalents of r_1 and r_2 . Thus, expected utility is transferable in this model. ■

Therefore, a sufficient condition for unique positive-assortative matching is supermodularity of the pairwise sum of certainty-equivalents, and a sufficient condition for unique negative-assortative matching is submodularity of the pairwise sum:

$$\begin{aligned}
CE(r_1, r_2) &= \gamma \hat{a}_{12} - c(\hat{a}_{12}) \\
&\quad - \frac{1}{r_1} \log \left(\frac{1}{2} \left[\frac{1}{\left[1 + \frac{r_2 c'(\hat{a}_{12}) V}{\gamma}\right]^{\frac{r_1}{r_2}} \left[1 - \frac{r_1 r_2}{r_1 + r_2} V\right]^{1 + \frac{r_1}{r_2}}} + \frac{1}{\left[1 - \frac{r_2 c'(\hat{a}_{12}) V}{\gamma}\right]^{\frac{r_1}{r_2}} \left[1 + \frac{r_1 r_2}{r_1 + r_2} V\right]^{1 + \frac{r_1}{r_2}}} \right] \right)
\end{aligned}$$

The first two terms of this sum can be thought of as the part of the expected utility pie coming from productivity (effort exertion and corresponding expected output), while the third term can

be thought of as the part of the expected utility pie coming from risk-sharing.

The challenge of identifying conditions for assortative matching in this model is the *one-sided* moral hazard. Although expected utility is transferable, r_1 and r_2 do not enter symmetrically into the sum of certainty-equivalents of the matched pair. Consequently, the matching conditions will *not* be distribution-free as they are in the absence of moral hazard (for example, in the first-best, or in the case of endogenous matching under the trade-off of *ex ante* and *ex post* risk management, as in Wang (2012c)).

Finding conditions for the supermodularity and submodularity of $CE(r_1, r_2)$ in r_1, r_2 yields the following matching results.

Proposition 16 *Let $c(a) = \eta a^M$, $M > 1$.*

1. *NAM is the unique eqm matching pattern for $M \in [1, M_1] \cup [M_4, \infty)$, where $M_1 \leq M_4$.*

a. *M_1 is increasing in r_1 (the least risk-averse principal's risk aversion) and in \bar{r}_1 (the most risk-averse principal's risk aversion), and decreasing in r_2, \bar{r}_2 . Furthermore, M_1 is increasing in V , and decreasing in γ .*

b. *M_4 is decreasing in r_1 and \bar{r}_1 , and increasing in r_2 and \bar{r}_2 . Furthermore, M_4 is decreasing in V , and increasing in γ .*

2. *PAM is the unique eqm matching pattern for $M \in [M_2, M_3]$, where $M_1 \leq M_2$ and $M_3 \leq M_4$.*

(It may be that the interval $[M_2, M_3]$ is empty.)

a. *M_2 is increasing in r_1 and \bar{r}_1 , and decreasing in r_2 and \bar{r}_2 . Furthermore, M_2 is increasing in V , and decreasing in γ .*

b. *M_3 is decreasing in r_1 and \bar{r}_1 , and increasing in r_2 and \bar{r}_2 . Furthermore, M_3 is decreasing in V , and increasing in γ .*

(The proof is relegated to the Appendix.)

This result highlights the key determinants of the equilibrium matching pattern: the riskiness of the environment V , the across-group and within-group heterogeneity in risk attitude, captured by the endpoints of the supports of the risk type distributions of principals and agents, and the marginal impact of effort on mean output, γ .

In words, the takeaways from the main matching result are the following. First, positive-assortative matching (PAM) is more likely for moderately convex cost of effort functions, while negative-assortative matching (NAM) is more likely for cost of effort functions which are close to linear or extremely convex.

What delineates the boundaries of "close to linear", "moderately" convex, and "extremely" convex? The comparative statics on the bounds tell us the following:

1. Negative-assortative matching (NAM) is more likely to arise when principals are distinctly more risk-averse than agents, while positive-assortative matching (PAM) is more likely to arise when principals are distinctly less risk-averse than agents.
2. NAM is more likely to arise when the environment is very risky (V is large), while PAM is more likely to arise when the environment is very safe (V is low).
3. NAM is more likely to arise when the marginal benefit of effort (for mean output) is low (γ is small), while PAM is more likely to arise when the marginal benefit of effort is high (γ is large).

What is the intuition behind these comparative statics? Consider a relatively safe environment where the cost of effort function is moderately convex and the marginal benefit of effort is large. Effort exertion across different risk attitudes is most heterogeneous when the cost of effort function is moderately convex. Moreover, when the environment is relatively safe (that is, V is relatively low), the "need" for insurance is small and output is a fairly precise signal of effort. Hence, rewarding effort based on realized output is effective, less risk-averse agents are substantially cheaper to incentivize than more risk-averse agents, and more risk-averse agents are not willing to pay particularly high risk premia. All of these forces push the incentive provision effect to outweigh the insurance provision effect, which favors PAM over NAM. If in addition principals are less risk-averse than agents, then the tradeoff between incentive provision and insurance provision is particularly stark (since if a principal were more risk-averse than the agent, the principal would be happy to provide incentives for the agent, as this would be a method of self-insurance).

Hence, in this environment, the least risk-averse principal *experiences the biggest difference in utility* between being paired with the least risk-averse agent versus a more risk-averse agent. Thus, the least risk-averse principal will outbid the other principals for the least risk-averse agent, and once the least risk-averse principal and the least risk-averse agent are removed from the pool of candidates, the least risk-averse principal of those remaining will outbid the other principals for the least risk-averse agent remaining, and so on, and the equilibrium matching pattern will be positive-assortative.

On the other hand, when agents have close to linear or extremely convex cost of effort, the difference between effort exertion across agents of different risk types is small—either all the agents

exert very high effort, or all the agents exert very low effort. If the environment is also risky, that is, V is high, then individuals, especially more risk-averse individuals, will be willing to pay a high price for insurance. Moreover, output is a noisy signal of actual effort exertion. Hence, the insurance provision effect will tend to outweigh the incentive provision effect. If in addition principals are more risk-averse than agents, then the incentive provision is aligned with insurance provision: a more risk-averse principal prefers a less risk-averse agent, because the principal desires insurance for herself, and this will naturally provide incentives to the agent. This means that the most risk-averse principal *experiences the biggest difference in utility* between being paired with the least risk-averse agent versus a more risk-averse agent. Thus, the most risk-averse principal will outbid the other principals for the least risk-averse agent, and once they are matched, the most risk-averse principal remaining will outbid the others for the most risk-averse agent remaining, and so forth, and the equilibrium matching pattern will be negative-assortative.

An interesting but informal insight that emerges from the analysis in this framework is that, in contrast with the standard view, there is a strong case for the more risk-averse individuals to be principals, and the less risk-averse individuals to be agents: a more risk-averse principal is happy to incentivize a less risk-averse agent, because she wants insurance, and the less risk-averse agent doesn't mind the riskiness of the incentives.

Finally, a word on efficiency.

Proposition 17 *The equilibrium maximizes the sum of certainty-equivalents, and is Pareto efficient.*

The equilibrium maximizes the sum of certainty-equivalents, since the conditions for PAM and NAM were derived by finding conditions for the supermodularity and submodularity of the pairwise sum of certainty-equivalents. Since the sum of certainty-equivalents is a social welfare function, and the equilibrium maximizes this sum, it must be Pareto efficient.

While the natural measure of welfare for a partnership and the individuals within that partnership is the pairwise certainty-equivalent, the unweighted sum of certainty-equivalents across pairs is not the right way of thinking about economy-wide welfare. A more risk-averse individual needs to be guaranteed a smaller amount than a less risk-averse individual to be made indifferent between accepting that amount with certainty and partaking in her risky equilibrium, but there is no reason society should value her less because of that. Hence, policy in this framework will be considered to improve aggregate welfare if it is Pareto-improving. (Alternatively, using a Rawlsian social wel-

fare function which weights the utility of more risk-averse individuals would also be a reasonable approach.)

2.4 The Policy Example

There are a variety of natural policies a government might wish to implement in this setting. In this hypothetical policy example, I will discuss a very simple approach to thinking about the welfare impacts of introducing formal insurance using this framework, and show how accounting for endogeneity is essential to the evaluation and design of this policy.

In particular, suppose that the status quo environment is very risky, that is, V is very large, and a government wishes to reduce the risk burden shouldered by risk-averse citizens. The government takes steps to reduce V by introducing formal insurance (modeling the introduction of formal insurance as a risk-reduction measure has precedent in the literature, for example in Attanasio and Rios-Rull (2000)). The matching results from Proposition 5 tell us that such a decrease in V may trigger an endogenous network response—in a very risky environment, the least risk-averse principals hire the most risk-averse agents, since a high V means that the gains from trade from risk-sharing are high, and moreover that output is a very noisy signal of effort. However, a decrease in V means that there is less risk that needs to be shared, and additionally that output is a much more precise signal of effort. Hence, the least risk-averse principals may switch to hiring the least risk-averse agents instead.

It will be helpful to work through a specific numeric example⁶. Let the parameters of the status quo be: $\gamma = 11.5$, $\eta = 0.5$, and $M = 4$, so that:

$$\begin{aligned} R &= 11.5a + \varepsilon, \varepsilon \sim \text{Laplace}(0, V) \\ c(a) &= \frac{1}{2}a^4 \end{aligned}$$

Suppose there are three principals and three agents: the principals are $\{r_1^1 = 0.3, r_2^1 = 0.8, r_3^1 = 1.5\}$, while the agents are $\{r_1^2 = 2, r_2^2 = 2.6, r_3^2 = 3\}$. That is, principals are distinctly less risk-averse than agents, and there is more heterogeneity in risk type amongst principals than amongst agents.

⁶The specific numbers are not important, as the results from Proposition 5 are comparative statics, and thus it is the relative comparisons that matter.

Suppose that initially, the level of risk in the economy is high. Specifically, the variance of returns of the project undertaken by all pairs is $V = 0.9$ (e.g., the terrain is such that rice paddy is the crop that all landlord-farmer pairs grow, and the standard deviation of profits from rice paddy is determined by V). Then the equilibrium match is negative-assortative, and the agent in each matched pair exerts the following effort:

$(r_3^1 = 1.5, r_1^2 = 2) : a_{r_1^2=2} = 1.34$
$(r_2^1 = 0.8, r_2^2 = 2.6) : a_{r_2^2=2.6} = 1.15$
$(r_1^1 = 0.3, r_3^2 = 3) : a_{r_3^2=3} = 0.87$

The least risk-averse agent exerts the highest level of effort, because she has an effectively lower marginal cost of effort, and, more importantly, works for the most risk-averse principal. Hence, the least risk-averse principal has very sharp incentives: she essentially insures the most risk-averse principal, and works hard herself, and this is a mutually-satisfying agreement.

The most risk-averse agent exerts the lowest level of effort, because she works for the least risk-averse principal, who provides her with informal insurance. The least risk-averse principal's land plot is therefore the least productive, but he is paid a risk premium by his agent.

The certainty-equivalents in each negatively-assorted partnership are:

$CE(r_3^1 = 1.5, r_1^2 = 2) = 13.08$
$CE(r_2^1 = 0.8, r_2^2 = 2.6) = 11.79$
$CE(r_1^1 = 0.3, r_3^2 = 3) = 9.53$

The happiest partnership is between the most risk-averse principal and the least risk-averse agent, because of the alignment between the principal's desire for insurance, and the agent's incentivization from a scheme which insures the principal. The unhappiest partnership is between the least risk-averse principal and the most risk-averse agent—even though there are gains from trade from risk-sharing, the loss of productivity resulting from the agent being insured reduces the joint expected utility pie.

Now suppose that the introduction of formal insurance reduces the riskiness of the environment, so that effectively the variance of the returns of the risky project falls to $V = 0.07$ (and the other parameters remain unchanged). This causes the principals and agents to re-sort: in this environment, the unique equilibrium match is positive-assortative in risk types.

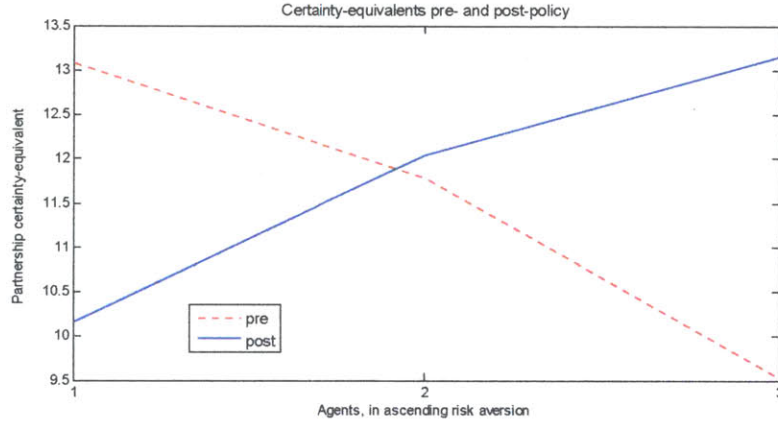
Then, the agent in each positively-assorted couple exerts effort:

$(r_1^1 = 0.3, r_1^2 = 2) : a_{r_1^2=2} = 0.914$
$(r_2^1 = 0.8, r_2^2 = 2.6) : a_{r_2^2=2.6} = 1.11$
$(r_3^1 = 1.5, r_3^2 = 3) : a_{r_3^2=3} = 1.25$

And the certainty-equivalent of each positively-assorted couple is:

$CE(r_1^1 = 0.3, r_1^2 = 2) = 10.16$
$CE(r_2^1 = 0.8, r_2^2 = 2.6) = 12.04$
$CE(r_3^1 = 1.5, r_3^2 = 3) = 13.15$

The certainty-equivalents of each partnership pre- and post-policy are plotted below. The agents are arranged along the x -axis in increasing risk aversion.



We can see a stark difference in the distribution of welfare pre- and post-policy. Pre-policy, in the risky environment, the least risk-averse agents are the best off, while they are the worst off following the introduction of formal insurance. Why is this?

Analysis of the impact of the introduction of formal insurance within this endogenous matching framework enables us to see precisely and quite concretely the often-discussed "crowding-out" effect. Following a decrease in V , insurance provision plays much less of a role in matching than does incentive provision, leading principals to hire agents with similar risk attitudes rather than different risk attitudes. This means that less insurance is provided informally in equilibrium.

Moreover, when V was high, the least risk-averse agents essentially insured the most risk-averse principals: the most risk-averse principals would offer their agents a wage scheme with high-powered incentives. This meant that the principal's income depended very little on the realized return of output, while the agent's income depended highly on realized return. But the agent, being less

risk-averse, was happy under this wage scheme and worked hard. This benefited both the principal and the agent: the principal was well-insured and expected return was high, while the agent was close to being a residual claimant to her effort exertion.

Following the introduction of formal insurance, however, the most risk-averse principal's desire for informal insurance is heavily dampened. She now prefers to take more risk, in the sense that she wants her own income stream to depend more on realized return of output. She therefore prefers to hire an agent who is also quite risk-averse.

Thus, the less risk-averse agent's role in providing informal insurance is "crowded out" by the introduction of formal insurance, and he is worse off, despite the drastic decrease in riskiness of the environment.

Although the approach to modeling the introduction of formal insurance is certainly oversimplistic in this example, the results are striking and capture the much-discussed "crowding out effect" concretely and rigorously. This analysis provides a strong argument for rigorously accounting for the impact on informal insurance relationships when introducing formal insurance.

2.5 Conclusion

A large literature has explored the consequences of balancing incentive provision with insurance provision for the contractual arrangement between a given principal and agent who work together to produce some output, where output depends on unobservable and noncontractible inputs put in by the agent. In this paper, I focus on the formation of the contracting relationships themselves. I argue that individuals in developing economies are typically forced to use their interpersonal relationships, which may already be serving several purposes, to address in addition the needs typically met by formal institutions in developed economies, such as risk management. I think of the equilibrium network of contracting relationships arising in this way as informal insurance.

I find conditions on the environment for assortative matching, in particular, on the level of riskiness in the environment, the marginal benefit of effort for mean output, the convexity of the cost of effort function, and the risk type distributions of principals and agents. I show that environments where principals are distinctly more risk-averse than agents, with high levels of risk, with low marginal benefit of effort, and with cost of effort functions which are close to linear or extremely convex are amenable to unique negative-assortative matching, while environments where principals are distinctly less risk-averse than agents, with low levels of risk, with high marginal

benefit of effort, and with cost of effort functions which are moderately convex are more conducive to positive-assortative matching. Loosely, the tradeoff that drives the equilibrium matching pattern is the costs and benefits of insurance provision versus the costs and benefits of insurance provision, for each possible pairing of risk types.

I then discuss applications of this analysis for policymaking. First, proper evaluation of policies which affect the parameters of this environment—for example, a policy which subsidizes inputs through an innovative technology and reduces the convexity of the cost of effort function, or a policy which introduces formal insurance and reduces the aggregate risk of the environment—requires accounting for the endogenous network response. A key point is that the multidimensionality of the relationships in the network is a consequence of missing formal institution, and hence partnerships between principals and agents, which are ostensibly formed for productivity purposes, may respond to changes in the risk environment, because embedded implicitly into these relationships is informal insurance. In particular, I showed that the introduction of formal insurance crowds out informal insurance, and may leave those individuals who acted as informal insurers worse off.

Second, this analysis is important for making accurate inferences about the environment from observations of equilibrium contracts. In particular, this analysis highlights the importance of collecting data about contracting partners, not just the sharing rules.

Much work remains to be done. A natural next step would be to allow principal-agent pairs to choose the riskiness of the income stream they face, instead of keeping the riskiness of the environment exogenous and the same for all possible pairs. It would also be interesting to allow for richer patterns of group formation. For example, tenant farmers working for different landlords might agree to share risk as a group. How would this affect the equilibrium network of relationships and contracts?

This analysis also yielded an insight into why more risk-averse individuals may be better suited as principals, while less risk-averse individuals may be better suited as agents, contrary to the standard perspective. A more risk-averse principal and a less risk-averse agent can be seen to work well together in a model of endogenous matching under risk and moral hazard, because the more risk-averse principal insures herself by providing incentives to her less risk-averse employee. Hence, in these relationships, insurance provision seems aligned with incentive provision. This suggests a closer study of risk attitudes and principal-agent roles in developing economies.

Thinking about informal institutions as the interpersonal relationships which emerge in equilibrium to address the needs usually served by formal institutions in developed economies gives us

a deeper understanding of the operation of informal institutions in the status quo, and how they would be affected by changes in the formal institutional structure. Furthering this understanding is essential for the design of effective development policies.

2.A Appendix

2.A.1 The one-sided CARA-Laplace moral hazard model

In this Appendix, I construct a model of one-sided moral hazard where both the principal and the agent are risk-averse (and may differ in the degree of their risk aversion). The principal and the agent have CARA utility, and the productive output of the pair depends noisily on the effort exerted by the agent, which is costly for the agent. The novelty is that the noise follows a Laplace distribution. I show that the unique equilibrium wage scheme in this model is piecewise linear. This is in contrast to the typical usage of a CARA-Normal framework with a linear wage justified by Holmstrom and Milgrom (1987) reasoning to counter the difficulty of characterizing the equilibrium of one-sided moral hazard models with risk-averse principals and agents, and a cumulative distribution function of returns which isn't globally convex. (Recall the standard monotone likelihood ratio and convexity of distribution function conditions, and slight variants, from Mirrlees (1979), Rogerson (1983), and Jewitt (1988). Convexity of the cumulative distribution function is usually the most difficult condition to satisfy to justify the first-order approach.)

In addition to the value of constructing a model in which the unique equilibrium can be characterized, and in particular has the convenient form of piecewise linearity, this model is well-suited to answering the specific question posed by this paper. I show that the equilibrium wage scheme is linear away from the mean, with the optimal risk-sharing slope—this captures insurance provision. And, the jump at the mean captures incentive provision: the agent is compensated linearly in output, and gets a bonus if output exceeds a threshold, where that threshold is the expected output given (correctly) anticipated effort.

The Model

The framework consists of the following elements:

The principal and the agent: both the principal r_1 and the agent r_2 are risk-averse with CARA utility $u(x) = -e^{-rx}$, $r > 0$. The principal owns one unit of physical capital but has infinite marginal cost of effort, while the agent owns no physical capital, but has a finite marginal cost of effort. The cost of effort for all agents is $c(a)$, $c(a) > 0$, $c'(a) > 0$, $c''(a) > 0$. The agent has an exogenously-given outside option denoted by $-e^{-v}$.

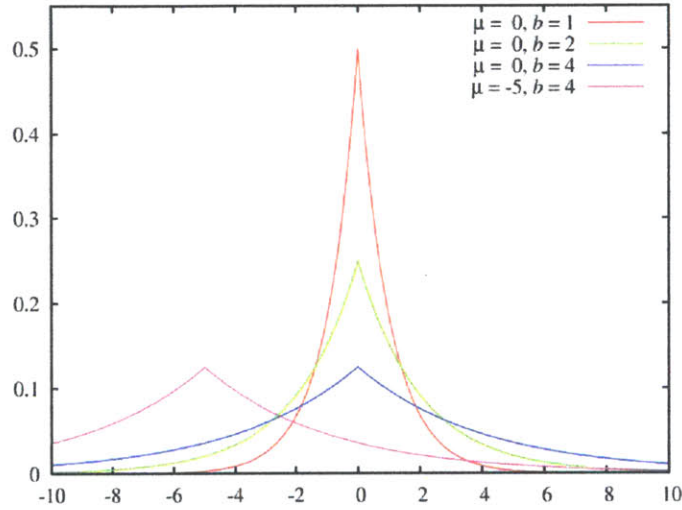
The risky environment: A principal-agent partnership can only produce positive output if one unit of physical capital is combined with human capital. Output is given by $R|a = \gamma a + \varepsilon$, where

the riskiness of the environment is captured by $\varepsilon \sim \text{Laplace}(0, V)$, which has support $(-\infty, \infty)$. Hence, the effort of an agent, a , increases the mean but doesn't affect the variance of returns.

The Laplace distribution is a continuous probability distribution with location and scale parameters $\mu \in \mathbb{R}$ and $b \in \mathbb{R}^+$, respectively. The pdf of a random variable $X \sim \text{Laplace}(\mu, b)$ is:

$$f(x|\mu, b) = \begin{cases} \frac{1}{2b} e^{-\frac{(\mu-x)}{b}}, & x < \mu \\ \frac{1}{2b} e^{-\frac{(x-\mu)}{b}}, & x \geq \mu \end{cases}$$

Pictorially, the pdf is:



Note that this resembles two back-to-back exponential distributions. In fact, if $Y \sim \text{Laplace}(0, V)$, then $|Y| \sim \exp(\frac{1}{V})$.

Information and commitment: The principal and the agent know each other's risk types. However, the agent's effort is not observable and not contractible.

A given principal and agent pair (r_1, r_2) observes the realized output of their partnership, and is able to commit to a return-contingent sharing rule $s(R)$, where R is realized output.

The equilibrium: An equilibrium consists of a wage $s(R)$ set by the principal satisfying the constraints of the problem, such that any other choice of feasible sharing rule by the principal would leave her worse off, as well as an effort level a chosen by the agent such that any other choice of effort would leave the agent worse off.

The Solution

The principal r_1 chooses wage $s(R)$ for the agent r_2 by solving the following problem:

$$\max_{s(R)} \int_{-\infty}^{\gamma a} -e^{-r_1(R-s(R))} \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR + \int_{\gamma a}^{\infty} -e^{-r_1(R-s(R))} \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR$$

such that:

$$\begin{aligned} (IR) : & \int_{-\infty}^{\gamma a} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR + \\ & + \int_{\gamma a}^{\infty} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR \geq -e^{-v} \\ (IC) : & a \in \arg \max_{a \in (0, \infty)} \int_{-\infty}^{\gamma a} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR + \\ & + \int_{\gamma a}^{\infty} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR \end{aligned}$$

It is clear that the *IR* constraint binds in equilibrium.

Since effort a is chosen from an open interval, all global maxima are stationary points, if they exist. Hence, the first-order problem (the problem wherein the global *IC* constraint is replaced by its first-order condition) is a relaxed problem: an optimum must be a stationary point, but a stationary point is not necessarily an optimum.

Differentiating the global *IC* constraint with respect to a yields the first-order condition of the global *IC* constraint:

$$\left[r_2 c'(a) - \frac{\gamma}{V} \right] \int_{-\infty}^{\gamma a} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR + \left[r_2 c'(a) + \frac{\gamma}{V} \right] \int_{\gamma a}^{\infty} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR = 0$$

Note that the compensation schedule $s(R)$ specifically *does not* depend on a , so differentiating the agent's expected utility from exerting effort a given $s(R)$ with respect to a does not require any assumptions about $s(R)$.

Solving the first-order problem yields the following wage schedule:

$$\begin{aligned} s(R < \gamma \hat{a}) &= \frac{r_1}{r_1 + r_2} R - \frac{r_1}{r_1 + r_2} \gamma \hat{a} + c(\hat{a}) + \frac{1}{r_2} v - \frac{1}{r_2} \log \left[\left(1 - \frac{r_1 r_2}{r_1 + r_2} V \right) \left(1 + \frac{r_2 c'(\hat{a}) V}{\gamma} \right) \right] \\ s(R > \gamma \hat{a}) &= \frac{r_1}{r_1 + r_2} R - \frac{r_1}{r_1 + r_2} \gamma \hat{a} + c(\hat{a}) + \frac{1}{r_2} v - \frac{1}{r_2} \log \left[\left(1 + \frac{r_1 r_2}{r_1 + r_2} V \right) \left(1 - \frac{r_2 c'(\hat{a}) V}{\gamma} \right) \right] \end{aligned}$$

where \hat{a} is the effort level "anticipated" by the principal.

Observe that if $\hat{a} = c'^{-1} \left(\frac{r_1}{r_1 + r_2} \gamma \right)$, then this compensation schedule is fully linear with slope

$\frac{r_1}{r_1+r_2}$; if $\hat{a} > c'^{-1}\left(\frac{r_1}{r_1+r_2}\gamma\right)$, then this compensation schedule is piecewise linear with a discrete jump at $\gamma\hat{a}$, and slope $\frac{r_1}{r_1+r_2}$ everywhere else, that is, at output levels $R \neq \gamma\hat{a}$.

What \hat{a} does the principal choose in equilibrium? Note that in equilibrium it must be that $a^* = \hat{a}$: that is, the principal would never pay for a level of effort higher than the one she anticipates, and an agent would never exert more effort than she is compensated for.

Now observe that for all $\hat{a} < \hat{a}_t$, for some threshold $\hat{a}_t > 0$, the unique stationary point of the agent r_2 's expected utility given effort exertion a and compensation scheme $s(R|\hat{a})$ is $a = \hat{a}$. Moreover, r_2 's expected utility given a and $s(R|\hat{a})$ is strictly concave in a , when $\hat{a} < \hat{a}_t$.

However, once $\hat{a} > \hat{a}_t$, where again, this \hat{a}_t is a threshold which will be rigorously characterized shortly, there are two stationary points of r_2 's expected utility given effort exertion a and compensation scheme $s(R|\hat{a})$: $a_1 = \hat{a}$ continues to be a stationary point, but $a_2 < \hat{a}$ is also a stationary point, and it can be seen that a_2 is the unique maximizer. The intuition is that, if the principal tries to induce "too much" effort, the agent will find it profitable to discretely deviate downwards.

Therefore, the equilibrium \hat{a} is $\hat{a} = \hat{a}_t$, where \hat{a}_t is the value of \hat{a} such that the second derivative of r_2 's expected utility given effort exertion a and compensation scheme $s(R|\hat{a})$ at $a = \hat{a}$ is precisely 0.

This yields the following expression characterizing $\hat{a}^* = \hat{a}_t$:

$$\left(c'(\hat{a}_t) - \frac{r_1}{r_1+r_2}\gamma\right) \left(c'(\hat{a}_t) + \frac{\gamma}{r_2V}\right) = \frac{1}{r_2}c''(\hat{a}_t) > 0$$

where it can clearly be seen that $\hat{a}_t > c'^{-1}\left(\frac{r_1}{r_1+r_2}\gamma\right)$, since $c(a)$ is increasing and convex, and the left-hand side is increasing in a .

But we know the agent's expected utility from exerting effort a given compensation schedule $s(R|\hat{a}^*)$ is strictly concave in a .

Therefore, the unique solution we found to the first-order problem is indeed the unique optimum.

This characterizes the equilibrium.

Incentive Provision and Insurance Provision

The piecewise linearity of the equilibrium wage schedule which emerges in this framework is a nice property for two key reasons. First, it's both tractable and realistic. Second, it neatly separates incentive provision from insurance provision: the linearity with slope $\frac{r_1}{r_1+r_2}$ at output levels above and below the mean captures efficient risk-sharing, while the discrete jump at the mean captures

incentive provision. Because the likelihood ratio is a piecewise constant with discontinuity at the mean output level, the key information about effort contained in output realization is whether the realized output is above or below $R = \gamma\hat{a}$. Conditional on knowing that realized output is below (above) mean output, however, no output level below (above) the mean is more informative about effort than another output level below (above) the mean. Furthermore, because the likelihood ratio assumes two "symmetric" values, $-Q$ and Q , the slope is the same for output levels above and below the mean.

To see more clearly how the piecewise linear wage scheme of the second-best solution cleanly separates incentive and insurance provision, it will be helpful to solve two cases: (a) the contracting principal between a risk-neutral principal and a risk-averse agent, where effort is not contractible, and (b) the contracting problem between a risk-averse principal and a risk-averse agent, where effort is contractible. The solution to the first case will be a step function, with one fixed wage for low output and a discretely higher fixed wage for high output, and the solution to the second case will be a perfectly linear wage with slope $\frac{r_1}{r_1+r_2}$, to capture efficient risk-sharing.

Case (a): the contracting problem between a risk-neutral principal and a risk-averse agent; effort not contractible The principal has utility $u(x) = x$, while the agent has utility $u(x) = -e^{-r_2x}$. Effort is not contractible. Hence, the principal chooses wage $s(R)$ for the agent r_2 by solving the following problem:

$$\max_{s(R)} \int_{-\infty}^{\gamma a} (R - s(R)) \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR + \int_{\gamma a}^{\infty} (R - s(R)) \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR$$

such that:

$$\begin{aligned} (IR) : & \int_{-\infty}^{\gamma a} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR + \\ & + \int_{\gamma a}^{\infty} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR \geq -e^{-v} \\ (IC) : & a \in \arg \max_{a \in (0, \infty)} \int_{-\infty}^{\gamma a} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR + \\ & + \int_{\gamma a}^{\infty} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR \end{aligned}$$

Replacing the global IC constraint with the first-order condition and writing the Lagrangean (letting λ be the multiplier on the binding IR constraint, and μ be the multiplier on the binding first-order condition):

$$\begin{aligned}
& \int_{-\infty}^{\gamma a} (R - s(R)) \frac{1}{2V} e^{\frac{1}{V}[R - \gamma a]} dR + \int_{\gamma a}^{\infty} (R - s(R)) \frac{1}{2V} e^{-\frac{1}{V}[R - \gamma a]} dR + \\
& + \lambda \left[\int_{-\infty}^{\gamma a} -e^{-r_2[s(R) - c(a)]} \frac{1}{2V} e^{\frac{1}{V}[R - \gamma a]} dR + \int_{\gamma a}^{\infty} -e^{-r_2[s(R) - c(a)]} \frac{1}{2V} e^{-\frac{1}{V}[R - \gamma a]} dR + e^{-v} \right] \\
& + \mu \left[\left[r_2 c'(a) - \frac{\gamma}{V} \right] \int_{-\infty}^{\gamma a} -e^{-r_2[s(R) - c(a)]} \frac{1}{2V} e^{\frac{1}{V}[R - \gamma a]} dR + \left[r_2 c'(a) + \frac{\gamma}{V} \right] \int_{\gamma a}^{\infty} -e^{-r_2[s(R) - c(a)]} \frac{1}{2V} e^{-\frac{1}{V}[R - \gamma a]} dR \right]
\end{aligned}$$

Then differentiating pointwise with respect to $s(R)$ for $R < \gamma a$ yields:

$$\begin{aligned}
-1 + \lambda r_2 e^{-r_2[s(R) - c(a)]} + \mu \left[r_2 c'(a) - \frac{\gamma}{V} \right] r_2 e^{-r_2[s(R_p) - c(a)]} &= 0 \Rightarrow \\
r_2 \left(\lambda + \mu \left[r_2 c'(a) - \frac{\gamma}{V} \right] \right) e^{-r_2[s(R_p) - c(a)]} &= 1 \\
\frac{1}{r_2} \log \left(r_2 \left(\lambda + \mu \left[r_2 c'(a) - \frac{\gamma}{V} \right] \right) \right) + c(a) &= s(R | R < \gamma a)
\end{aligned}$$

Hence, $s(R)$ for $R < \gamma a$ is:

$$s(R | R < \gamma a) = \frac{1}{r_2} \log \left(r_2 \left(\lambda + \mu \left[r_2 c'(a) - \frac{\gamma}{V} \right] \right) \right) + c(a)$$

And, differentiating pointwise with respect to $s(R)$ for $R > \gamma a$ yields:

$$s(R | R > \gamma a) = \frac{1}{r_2} \log \left(r_2 \left(\lambda + \mu \left[r_2 c'(a) + \frac{\gamma}{V} \right] \right) \right) + c(a)$$

Hence, the equilibrium wage contract for the agent r_2 is a wage fixed at a certain level for $R < \gamma \hat{a}$, and a wage fixed at a discretely higher level for $R > \gamma \hat{a}$, where \hat{a} is the effort level anticipated by the principal. So, the wage looks like, "flat, jump at the anticipated mean level of output, and flat again".

Case (b): the contracting problem between a risk-averse principal and a risk-averse agent; effort contractible Now, both the principal and the agent are risk-averse with CARA utility, but effort is contractible. The principal r_1 solves the following problem:

$$\max_{a, s(R)} \int_{-\infty}^{\gamma a} -e^{-r_1(R - s(R))} \frac{1}{2V} e^{\frac{1}{V}[R - \gamma a]} dR + \int_{\gamma a}^{\infty} -e^{-r_1(R - s(R))} \frac{1}{2V} e^{-\frac{1}{V}[R - \gamma a]} dR$$

such that:

$$(IR) : \int_{-\infty}^{\gamma a} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{\frac{1}{V}[R-\gamma a]} dR + \\ + \int_{\gamma a}^{\infty} -e^{-r_2[s(R)-c(a)]} \frac{1}{2V} e^{-\frac{1}{V}[R-\gamma a]} dR \geq -e^{-v}$$

Letting λ denote the Lagrange multiplier on the IR constraint, differentiating pointwise with respect to $s(R)$ for $R < \gamma a$ yields:

$$-r_1 e^{-r_1(R-s(R))} + \lambda r_2 e^{-r_2[s(R)-c(a)]} = 0$$

And differentiating pointwise with respect to $s(R)$ for $R > \gamma a$ yields:

$$-r_1 e^{-r_1(R-s(R))} + \lambda r_2 e^{-r_2[s(R)-c(a)]} = 0$$

the same condition.

Hence:

$$s(R|\hat{a}) = \frac{r_1}{r_1 + r_2} R + \frac{r_2}{r_1 + r_2} c(\hat{a}) \frac{1}{r_1 + r_2} \log \left(\lambda \frac{r_2}{r_1} \right)$$

a linear wage schedule with slope $\frac{r_1}{r_1 + r_2}$.

The key takeaway from both of these cases is that, in this model, it is possible to see the effect of both the principal and the agent being risk-averse on the equilibrium wage, *and* the effect of effort not being observable or contractible. The equilibrium wage schedule that emerges when both of these components are combined is the wage schedule which is linear at output levels away from the mean (as in Case (b)), and where there is a discontinuous jump at the mean (as in Case (a)).

2.A.2 The equilibrium for a given pair in the first-best

First, suppose principals have a more general increasing, strictly concave utility function given by $U(x; p)$, and agents have a more general increasing, strictly concave utility function given by $V(x; q)$.

Then, observe that the constraint set is convex. That is, for a given a , if $s_1(R)$ satisfies IR , and $s_2(R)$ satisfies IR , then $\alpha s_1(R) + (1 - \alpha)s_2(R)$ also satisfies IR .

If $s_1(R)$ satisfies IR , and $s_2(R)$ satisfies IR , then:

$$\begin{aligned}\int_{-\infty}^{\infty} V(s_1(R) - c(a))f(R - \gamma a)dR &\geq v \\ \int_{-\infty}^{\infty} V(s_2(R) - c(a))f(R - \gamma a)dR &\geq v\end{aligned}$$

Then: use the concavity of $V(\cdot)$:

$$\begin{aligned}\int_{-\infty}^{\infty} V(\alpha s_1(R) + (1 - \alpha)s_2(R) - c(a))f(R - \gamma a)dR &= \\ \int_{-\infty}^{\infty} V[\alpha(s_1(R) - c(a)) + (1 - \alpha)(s_2(R) - c(a))]f(R - \gamma a)dR &> \\ \int_{-\infty}^{\infty} \alpha V(s_1(R) - c(a))f(R - \gamma a)dR + \int_{-\infty}^{\infty} (1 - \alpha)V(s_2(R) - c(a))f(R - \gamma a)dR &\geq \\ \alpha v + (1 - \alpha)v &= v\end{aligned}$$

And, the objective function is strictly quasiconcave: the objective function is:

$$\int_{-\infty}^{\infty} U(R - s(R))f(R - \gamma a)dR$$

Fix a . Then, pointwise differentiate twice wrt $s(R)$ for each R :

$$\begin{aligned}f' &: -U'(R - s(R))f(R - \gamma a) \\ f'' &: U''(R - s(R))f(R - \gamma a) < 0\end{aligned}$$

since $U(\cdot)$ is concave.

So, there should be a unique global constrained maximizer.

Now, let's solve for the first-best equilibrium when principals and agents explicitly have CARA utility, and differ in their Arrow-Pratt coefficient of risk aversion, r .

The first-order conditions of the optimization problem are:

$$\begin{aligned}
\frac{r_1}{r_2} e^{-r_1 R + (r_1 + r_2)s(R) - r_2 c(a)} &= \lambda \text{ for each } R \\
\frac{c'(a)}{\gamma} \frac{\int_{-\infty}^{\infty} r_2 e^{-r_2[s(R) - c(a)]} f(R - \gamma a) dR_p}{\int_{-\infty}^{\infty} e^{-r_1[R - s(R)]} f'(R - \gamma a) dR_p} - \frac{\int_{-\infty}^{\infty} e^{-r_2[s(R) - c(a)]} f'(R - \gamma a) dR}{\int_{-\infty}^{\infty} e^{-r_1[R - s(R)]} f'(R - \gamma a) dR} &= \frac{1}{\lambda} \\
\int_{-\infty}^{\infty} -e^{-r_2[s(R) - c(a)]} f(R - \gamma a) dR &= -e^{-v}
\end{aligned}$$

We know that in the first-best, under the general $U(x; p), V(x; q)$ utility functions, the slope of the equilibrium wage will be:

$$s'_{FB}(R) = \frac{1}{1 + \left(-\frac{U'(R - s(R))}{U''(R - s(R))} \right) \left(-\frac{V''(s(R) - c(a))}{V'(s(R) - c(a))} \right)}$$

When the utility functions are CARA, this expression is equal to $\frac{r_1}{r_1 + r_2}$. So, under CARA utility, the equilibrium wage of the agent is linear with slope $\frac{r_1}{r_1 + r_2}$: this means that $s^*_{FB}(R_p) = \frac{r_1}{r_1 + r_2} R_p + K$, where K is some constant depending on parameters. This is intuitive: effort is contractible, so there is no need to provide incentives. Hence the sharing rule efficiently shares risk; this means the schedule is linear. Moreover, the less risk-averse individual will have the share that is more dependent on output realization.

Thus, our conditions are:

$$\begin{aligned}
\frac{r_1}{r_2} e^{(r_1 + r_2)K - r_2 c(a)} &= \lambda \\
\frac{c'(a) r_2 e^{-(r_1 + r_2)K} e^{r_2 c(a)}}{\gamma} \frac{\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R} f(R - \gamma a) dR}{\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R} f'(R - \gamma a) dR} - e^{-(r_1 + r_2)K} e^{r_2 c(a)} &= \frac{1}{\lambda} \\
e^{-r_2(K - c(a))} \int_{-\infty}^{\infty} -e^{-\frac{r_1 r_2}{r_1 + r_2} R} f(R - \gamma a) dR &= -e^{-v}
\end{aligned}$$

Observe that we can integrate by parts:

$$\begin{aligned}
\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R} f'(R - \gamma a) dR &= e^{-\frac{r_1 r_2}{r_1 + r_2} R} f(R - \gamma a) \Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} \left(-\frac{r_1 r_2}{r_1 + r_2} \right) e^{-\frac{r_1 r_2}{r_1 + r_2} R} f(R - \gamma a) dR \\
&= \frac{r_1 r_2}{r_1 + r_2} \int_{-\infty}^{\infty} -e^{-\frac{r_1 r_2}{r_1 + r_2} R} f(R - \gamma a) dR
\end{aligned}$$

Hence:

$$\frac{\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R} f(R - \gamma a) dR}{\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R} f'(R - \gamma a) dR} = \frac{r_1 + r_2}{r_1 r_2}$$

Then the first two conditions imply:

$$\begin{aligned} \frac{r_2}{r_1} e^{-(r_1 + r_2)K + r_2 c(a)} &= \frac{c'(a) r_2 e^{-(r_1 + r_2)K} e^{r_2 c(a)}}{\gamma} \left(\frac{r_1 + r_2}{r_1 r_2} \right) - e^{-(r_1 + r_2)K} e^{r_2 c(a)} \Rightarrow \\ a_{FB}^* &= c'^{-1}(\gamma) \end{aligned}$$

Hence, first-best effort is independent of risk type. Any employer r_1 who hires an employee r_2 will require her to exert effort $a_{FB}^* = c'^{-1}(\gamma)$. That is, first-best effort equates marginal benefit of effort (marginal impact of effort on mean output) with marginal cost.

Finally, use K to satisfy r_2 's IR constraint:

$$\begin{aligned} e^{-r_2(K - c(c'^{-1}(\gamma)))} \int_{-\infty}^{\infty} -e^{-\frac{r_1 r_2}{r_1 + r_2} R} f(R - \gamma c'^{-1}(\gamma)) dR &= -e^{-v} \Rightarrow \\ K_{FB}^*(r_1, r_2, v) &= \frac{1}{r_2} v + c(c'^{-1}(\gamma)) + \frac{1}{r_2} \log \left(\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R} f(R - \gamma c'^{-1}(\gamma)) dR \right) \end{aligned}$$

Hence, the optimal first-best contract is:

$$\begin{aligned} s_{FB}^*(R) &= \frac{r_1}{r_1 + r_2} R + K_{FB}^*(r_1, r_2, v) \\ K_{FB}^*(r_1, r_2, v) &= \frac{1}{r_2} v + c(c'^{-1}(\gamma)) + \frac{1}{r_2} \log \left(\int_{-\infty}^{\infty} e^{-\frac{r_1 r_2}{r_1 + r_2} R} f(R - \gamma c'^{-1}(\gamma)) dR \right) \\ a_{FB}^* &= c'^{-1}(\gamma) \end{aligned}$$

2.A.3 Proof of the main result in the second-best

We know that the sum of certainty-equivalents in a pair (r_1, r_2) is:

$$CE(r_1, r_2) = \gamma \hat{a}_t - c(\hat{a}_t) - \frac{1}{r_1} \log \left(\frac{1}{2} \left[\frac{1}{\left[1 + \frac{r_2 c'(\hat{a}_t) V}{\gamma} \right]^{\frac{r_1}{r_2}} \left[1 - \frac{r_1 r_2}{r_1 + r_2} V \right]^{1 + \frac{r_1}{r_2}}} + \frac{1}{\left[1 - \frac{r_2 c'(\hat{a}_t) V}{\gamma} \right]^{\frac{r_1}{r_2}} \left[1 + \frac{r_1 r_2}{r_1 + r_2} V \right]^{1 + \frac{r_1}{r_2}}} \right] \right)$$

where \hat{a}_t is characterized by:

$$\left(c'(\hat{a}_t) - \frac{r_1}{r_1 + r_2} \gamma \right) \left(c'(\hat{a}_t) + \frac{\gamma}{r_2 V} \right) = \frac{1}{r_2} c''(\hat{a}_t) > 0$$

We can easily see that $\hat{a}_t > c'^{-1} \left(\frac{r_1}{r_1 + r_2} \gamma \right)$.

Replacing the functional form for cost of effort into the expressions:

$$\begin{aligned} c(a) &= \eta a^M, M > 1 \\ c'(a) &= \eta M a^{M-1} \\ c''(a) &= \eta M(M-1) a^{M-2} \end{aligned}$$

Then the expressions characterizing the pairwise sum and the threshold \hat{a}_t become:

$$CE(r_1, r_2) = \gamma \hat{a}_t - c(\hat{a}_t) - \frac{1}{r_1} \log \left(\frac{1}{2} \left[\frac{1}{\left[1 + \frac{r_2 \eta M \hat{a}_t^{M-1} V}{\gamma} \right]^{\frac{r_1}{r_2}} \left[1 - \frac{r_1 r_2}{r_1 + r_2} V \right]^{1 + \frac{r_1}{r_2}}} + \frac{1}{\left[1 - \frac{r_2 \eta M \hat{a}_t^{M-1} V}{\gamma} \right]^{\frac{r_1}{r_2}} \left[1 + \frac{r_1 r_2}{r_1 + r_2} V \right]^{1 + \frac{r_1}{r_2}}} \right] \right)$$

and:

$$\left(\eta M \hat{a}_t^{M-1} - \frac{r_1}{r_1 + r_2} \gamma \right) \left(\eta M \hat{a}_t^{M-1} + \frac{\gamma}{r_2 V} \right) = \frac{1}{r_2} \eta M(M-1) \hat{a}_t^{M-2}$$

Finding the conditions for supermodularity and submodularity produces the comparative statistics.

Chapter 3

Interdependent Utility and Truthtelling in Two-Sided Matching

3.1 Introduction

Results from the economic study of market design have been used quite successfully in a variety of "real-world" markets in which price mechanisms fail or are not available. *Stability* was established early on as a natural equilibrium concept—in a world where agents belong to one of two groups, and gain by working together across groups, a one-to-one matching is stable if it satisfies two properties: first, each matched agent must prefer being with its partner to remaining single, and second, no two agents who are unmatched in the assignment can match with each other instead and both become better off. Gale and Shapley (1962) proposed the simple but powerful deferred acceptance algorithm (DAA), which arrives at a stable matching given the ordinal preferences reported by each agent, and is used to facilitate a number of real-life two-sided matching markets, including the assignment of students to charter schools, and the assignment of medical students to residencies. Since it implements a stable matching, we call the DAA a stable mechanism. According to the DAA, one side is chosen to "propose". For example, suppose that in a standard marriage problem, the men are chosen to propose to the women. Then, each man proposes to his first choice. If a woman receives a single proposal, she holds it. If a woman receives multiple proposals, she holds her most preferred one and rejects the others. Men who were rejected then propose to their next favorite choice, and so on. A man never re-proposes to a woman who rejected him. The stable matching selected is called the man-optimal stable matching; when women propose, that stable

matching is woman-optimal. The widespread usage of the DAA is due not only to its simplicity, but also to the fact that it is observed to work quite well in practice. Kojima and Pathak (1996) remark, "In real-world applications, empirical studies have shown that stable mechanisms often succeed, whereas unstable ones often fail." Roth (1989) provides a large body of evidence.

Yet the theoretical manipulability of stable mechanisms is well-known. For example, Roth (1982) shows that there is no stable mechanism where truthfully reporting preferences is a dominant strategy for every agent. In addition, Roth and Sotomayor (1990) show that, when any stable mechanism is applied to a marriage market in which preferences are strict and there is more than one stable matching, then at least one agent can misreport her preferences and become matched to a partner more desirable than the one she would have been assigned had she reported the truth, assuming all the other agents are telling the truth. Thus, in situations where there are multiple stable matchings with respect to the set of genuine preferences, at least one person always has an incentive to lie. Moreover, loosely speaking, there are often multiple stable matchings—the core is always nonempty, and the stable matching arrived at by the DAA when men propose generally differs from the matching arrived at when the women propose. Importantly, there are some limits to manipulability: Dubins and Freedman (1981) and Roth (1982) showed that under the man-optimal DAA, it is a dominant strategy for every man to report preferences truthfully, and similarly for the woman-optimal DAA and women. However, women still have incentives to misreport under the man-optimal DAA, and vice versa.

What might reconcile this apparent paradox of manipulation in theory but not in practice? The aim of this paper is to offer one explanation not yet explored in the literature for why agents do not seem to take advantage of opportunities to misreport profitably. The intuition is that an agent's utility from a partnership may depend not only on her own views about her partner, but also on her partner's views of her. For example, in the student-school matching problem, a student considers the public profile of characteristics for each of the schools: student-teacher ratio, extracurriculars, financial aid, facilities, and so on. Based on these public profiles, the students rank the schools.

However, the students may very well care about how those schools rank them. If the student thinks a school is great because it has small classes and well-known teachers, but discovers that the school doesn't think very highly of the student, the student might re-think ranking that school highly. After all, if she ends up at the school, she may not receive any attention or resources, making it not such a great option after all.

This interdependence of utility is a realistic feature of many other matching scenarios, including

employment and dating/marriage. Working at one's dream firm is not enjoyable if one feels overlooked or undervalued; dating a dream person is not pleasant if that person is dissatisfied. Hence, an agent may care about her desires as well as her desirability in a given match.

I formalize this notion of interdependent utility and explore stable matchings in this framework. Suppose that each market participant has a publicly observable profile of characteristics. The participants care about these characteristics, because once matched, a partnership can produce valued output, and production depends positively on the effort exerted by each partner. The amount of effort a partner exerts depends on her happiness with the partnership. After observing the public profiles of potential partners, an agent is able to determine how satisfied she would be in a relationship with each of the potential partners, and therefore how much effort she would exert in each relationship—she exerts more effort in relationships with people who have characteristics she likes. Thus, she forms "first-round" ordinal preferences over her potential partners. However, she doesn't know how much effort a potential partner would exert in the relationship without knowing something about how satisfied that person is with being her partner. An indicator of the level of satisfaction is that partner's ranking of her. Hence, in this formulation, an agent's "final" utility from a partnership depends on her "first-round" ranking of her partner as well as her partner's "first-round" ranking of her, because the output the partnership ultimately produces depends on both partners' contributions, and an agent's contribution is increasing in her satisfaction.

I show that, for sets of preferences where, in the standard model, there are multiple stable matchings and at least one person has an incentive to misreport, in this framework, there is either a unique stable matching and no agent has an incentive to lie given that everybody else is telling the truth, or there are multiple stable matchings, but still no agent has an incentive to lie given that everybody else is telling the truth. (By "the standard model", I mean the traditional one-to-one, two-sided matching model where each agent submits ordinal preferences, and the utility an agent i gets from being matched with j is higher the more highly she ranks j .) The intuition for this is the following: in the standard model, an agent might misreport her preferences (e.g. declare truncated preferences) in the hopes of getting partnered with someone she has ranked more highly than the person she would be assigned if she reported her preferences truthfully. However, the reason she isn't assigned to the higher-ranked person if she reports truthfully is because that person isn't as enthusiastic about being partnered with her. In the standard, "independent utility" model, she doesn't care if her partner is enthusiastic about her or not—she just wants to end up with someone as high up her list as possible. But in the interdependent utility model of this paper, our agent cares

how enthusiastic her partner is about her. She doesn't find it worthwhile to lie about preferences in order to be partnered with someone whom she ranks very highly, but who ranks her very poorly.

Put another way, in the standard model, multiple stable matchings are associated with across-group disagreement. If, for each school, a school's most preferred student also ranks that school as most preferred, then there will be a unique stable matching. Multiplicity arises if, for instance, the school's most preferred student ranks a different school as her favorite. Interdependent utility "smooths" this across-group disagreement by reducing the distance, so to speak, between two agents' feelings about each other. In the standard model, an agent i in one group could despise agent j in the other group (rank him last), while agent j could adore agent i (rank her first). Hence, the stable match that the DAA selects when one group proposes may differ drastically from the stable match that the DAA selects when the other group proposes. Under interdependent utility, agent i 's disapproval of j is tempered by the knowledge that j is devoted to i , and j 's adoration of i is tempered by the knowledge that he is i 's last choice. This diminished disparity in preferences across groups pushes the market participants towards a sort of agreement on a unique stable matching. In the limiting case where each agent equally weights her desires and her desirability, there is in fact a unique stable matching in almost every instance of ordinal preferences. Moreover, this unique stable matching is either the man-optimal or the woman-optimal matching in the standard framework. In addition, I identify a set of ordinal preferences which I call "perfectly antagonistic"—given these preferences, even under interdependent utility, there is not a unique stable matching (disagreement is irreconcilable, in some sense), but nevertheless it continues to be the case no agent has an incentive to misreport preferences.

Having shown that, in contrast with the standard model, the stable match is generally unique in a framework with interdependent utility where desires and desirability are equally weighted, and hence agents do not have an incentive to misreport their preferences, I then construct an appropriately-modified DAA, based off the sum of the ranks in each potential partnership between agent i in group one and agent j in group two, which selects the unique stable match. To be a bit more specific, a matrix is constructed, where the $(i, j)^{th}$ entry is the sum of the i^{th} member of group one's ranking of the j^{th} member of group two, and the j^{th} member of group two's ranking of the i^{th} member of group one. Since the stable match is unique, it doesn't matter which side proposes. Suppose WLOG that group one proposes. Then, each member of group one proposes to the member of group two with whom she has the minimal rank-sum. Ties can be broken randomly, or by using the initial ordinal rankings of the proposer. If a member of group two receives a single

proposal, she holds it; if she receives multiple proposals, she keeps the one with minimal rank-sum and rejects the others. Any member of group one who was rejected then proposes to the member of group two with whom she has the next smallest rank-sum, and so on. No member ever re-reproposes to someone who already rejected her.¹

Other explanations exist in the literature for why misreporting does not seem to occur under stable mechanisms. One branch studies the computational complexity of calculating the optimal lie. Some researchers have used methods from computer science to show that the problem of profitable manipulation may be NP hard (Pini et al. 2009). On a related note, Roth and Rothblum (2002) argue that part of the difficulty of constructing the optimal lie may stem from incomplete information. Perhaps one needs to know the preferences of one's fellow group members, or the preferences of the members of the other group. In fact, Roth and Rothblum show that, even when information is highly incomplete, reporting truncated preferences (that is, declaring one's least preferred options to be unacceptable when they are actually acceptable, just heavily disliked) is still profitable. Truncation strategies are also not particularly computationally complex.

Another set of papers argues that there is little to gain from manipulation in large markets. Kojima and Pathak (2009) show that in many-to-one matching markets, such as students to schools, as the number of participants goes to infinity but the length of preference lists stays fixed, the fraction of participants with an incentive to misreport their preferences, given that everyone else is telling the truth, goes to 0. More specifically, under the student-optimal stable mechanism in the schools-students matching problem, we know that it is a dominant strategy for students to report their preferences truthfully, but that schools have an incentive to misreport their quotas. Kojima and Pathak show that the benefit to schools from manipulating their quotas diminishes as the market grows large.

The key distinction between the existing explanations for the absence of manipulation of stable mechanisms in practice and the explanation advanced by this paper is that in existing frameworks, market participants *do* have an incentive to misreport their preferences, only it is difficult to figure out how, or the gain from it is small. By contrast, the results of this paper suggest that if an agent simply cares not only about being matched with a partner she likes, but also about being matched with a partner who likes her, then in fact there will be a unique stable matching and no agent will

¹The model, the algorithm (including how to deal with declaring unacceptable agents), and the results will be presented much more precisely in the coming sections. This is merely meant to serve as an overview of the main idea of the paper.

have an incentive to lie if the others are telling the truth.

The rest of the paper proceeds as follows. In the next section, I work through a simple example to demonstrate how an agent may gain in the standard framework by misreporting her preferences. I then set up the model of interdependent utility, and present the results. I revisit the example to demonstrate that agents no longer have the incentive to lie in a framework with interdependent utility. Finally, I conclude.

3.2 An Example

Consider a standard marriage market with four women and four men. Suppose nobody prefers being single to being matched, and ordinal preferences are:

$$M1 : W2 > W3 > W1 > W4$$

$$M2 : W1 > W2 > W4 > W3$$

$$M3 : W4 > W3 > W2 > W1$$

$$M4 : W1 > W3 > W2 > W4$$

$$W1 : M1 > M3 > M2 > M4$$

$$W2 : M1 > M3 > M4 > M2$$

$$W3 : M2 > M1 > M3 > M4$$

$$W4 : M4 > M2 > M3 > M1$$

Then, in the first round of the man-proposing DAA, the men propose as follows:

$$M1 \rightarrow W2$$

$$M2 \rightarrow W1$$

$$M3 \rightarrow W4$$

$$M4 \rightarrow W1$$

Since $W1$ receives two proposals, she holds the one she prefers, which is $M2$'s proposal. Thus, $M4$ must propose to his next favorite woman, $W3$, who has no competing proposals. The DAA ends at this point, and the man-optimal stable matching is:

$$\mu_M = \begin{bmatrix} M1 & M2 & M3 & M4 \\ W2 & W1 & W4 & W3 \end{bmatrix}$$

The woman-proposing DAA runs in a similar fashion, except the women propose. This produces the woman-optimal stable matching:

$$\mu_W = \begin{bmatrix} M1 & M2 & M3 & M4 \\ W2 & W3 & W1 & W4 \end{bmatrix}$$

Note that $\mu_M \neq \mu_W$, so there are multiple stable matchings. We know from Dubins and Freedman [4] and Roth [10] that given these preferences, under any stable mechanism, at least one person has an incentive to misreport preferences given the others are telling the truth.

For example, suppose the man-proposing DAA is being run, and $W4$ "misreports" her preferences by truncating them. That is, $W4$ reports that the only man acceptable to her is $M4$ (but her true preferences are as described above).

Then, in the first round of the man-proposing DAA, the men propose as follows:

$$M1 \rightarrow W2$$

$$M2 \rightarrow W1$$

$$M3 \rightarrow W4$$

$$M4 \rightarrow W1$$

$W1$ again receives multiple proposals from $M2$ and $M4$, and holds the one she prefers, $M2$'s, but in addition $W4$ rejects $M3$'s proposal, because she has declared him to be unacceptable. Hence, $M4$ and $M3$ have to propose again. $M4$ proposes to his next favorite woman, $W3$, while $M3$ proposes to $W3$ as well.

$$M1 \rightarrow W2$$

$$M2 \rightarrow W1$$

$$M3 \rightarrow W3$$

$$M4 \rightarrow W3$$

$W3$ holds the proposal she prefers, $M3$'s, so $M4$ proposes to $W2$.

$$M1 \rightarrow W2$$

$$M2 \rightarrow W1$$

$$M3 \rightarrow W3$$

$$M4 \rightarrow W2$$

But $W2$ already has a proposal from $M1$, who is her top choice, so she rejects $M4$, who finally proposes to $W4$, and she accepts. This leads to the matching:

$$\begin{bmatrix} M1 & M2 & M3 & M4 \\ W2 & W1 & W3 & W4 \end{bmatrix}$$

So, by lying and truncating her preferences, $W4$ ends up with her top choice of partner, $M4$, under the man-proposing DAA, whereas she would have ended up with $M3$ if she had told the truth. In this model, an agent is strictly happier to be paired with a partner she ranks more highly.

Now, let's turn to the model of this paper.

3.3 The Model

The economy is composed of two disjoint groups of rational, utility-maximizing agents, who have utility functions $u(x) = x$. Denote the groups by M and W , where $|M| = |W| \in \{2, 3, \dots, N\}$, $N < \infty$. Suppose that members of M and W have to build partnerships one-to-one across groups in order to produce some kind of valuable output. For example, these two groups could be men and women, or students and schools, or employees and employers.

Suppose that each agent i has a public profile of characteristics, $X_i \in \mathbb{R}^k$, $k \in \mathbb{N}$, which is known to herself and observable to others. For example, for a school, this public profile might consist of the average student-teacher ratio, the number of classrooms, the extracurriculars available, the facilities, teacher quality, and so forth. For a student, this public profile might consist of past grades and past exam scores.

Suppose that the output produced by a matched partnership (i, j) depends on the characteristics of the partners, X_i and X_j , in the following way. Any matched partnership (i, j) has the possibility of capturing a pie of size y . However, once matched, i and j must contribute to the relationship in order to capture some fraction of the pie. The level of an agent's contribution depends on her compatibility and satisfaction with her partner. Since a contribution is personally costly, an agent will contribute highly only if she is satisfied with the relationship. The joint contributions of the agents determine the output produced; output is increasing in the level of contribution.

In particular, an agent $i \in M$ matched with an agent $j \in W$ chooses a contribution:

$$c_i^j : X_i \times X_j \rightarrow \mathbb{R}$$

(There is no restriction or assumption on the form of this mapping, or on the profiles of characteristics.)

So, $\{c_i^1, \dots, c_i^N\}$ is the vector whose j^{th} element describes the contribution i would make in a relationship with the j^{th} member of W . Let $\rho(c_i^j)$ denote the ranking of the size of the contribution c_i^j relative to the other elements in the vector, from largest to smallest. For example, if c_i^j is the largest number in the vector, then $\rho(c_i^j) = 1$, indicating that i perceives j to be her most compatible partner, and is most willing to make contributions in a relationship with him. Assume that no elements of the vector are the same, so that each i has strict preferences over possible partners j , based on her evaluation of the compatibility of their characteristics.

Suppose that the fraction of the pie of size y captured by a partnership (i, j) is $\theta(i, j)$, where

$$\theta(i, j) = f(\rho(c_i^j)) + f(\rho(c_j^i))$$

and

$$f(x) = \frac{1}{2} \frac{N - x}{N - 1}$$

So, the output produced by a partnership (i, j) is:

$$\theta(i, j)Y = \frac{N - \left(\frac{\rho(c_i^j) + \rho(c_j^i)}{2} \right)}{N - 1} Y$$

where $\rho(\cdot) \in \{1, \dots, N\}$ since it is a ranking.

For instance, suppose that a student i has the characteristic that she wants to become a mathematician, and so wishes to match with a school which prioritizes the sciences and excels at math competitions. Suppose school j prioritizes the sciences, while school k prioritizes theater. Then, student i recognizes she is much more compatible with school j than with school k —this matters because she knows that at school j , she will work very hard and become involved with the math competitions, while at the same time receiving the support and the resources that she needs from school j , which is excited to have her as a student. By contrast, if she were to attend school k , the student knows she would not exert very much effort in theater, and would likely also be ignored by the school, as they prefer to devote their resources to those with talent in theater.

Hence, student i ranks school j above school k —not only does i prefer j to k because she prefers science to theater, she also knows that school j will give her time and resources. Hence, both i and j will contribute highly to a relationship with each other, and this will lead to high productivity. In particular, if i and j rank each other first, that is, $c_i^j = c_j^i = 1$, then $f(c_i^j) = f(c_j^i) = \frac{1}{2}$, so that $\theta(i, j) = 1$, and a partnership between student i and school j captures the entire pie y .

Note that in this model, an agent always prefers being matched to somebody rather than remaining single. However, when I discuss a mechanism to implement stable matches in this setting, I will assume that the "central computer" which is running the mechanism does not know that nobody prefers being single, and so agents are able to misreport partners as being unacceptable if they wish.

A match function μ assigns distinct members of M to distinct members of W , where each person is assigned at most one partner. A matching is an equilibrium if it is *stable*—that is, it must satisfy two properties:

1. Individual rationality: every agent prefers being matched with her partner to remaining single (this is satisfied for every agent in this model)
2. No blocking: no two agents who are unmatched under μ can make themselves both better off by matching with each other instead of obeying the assignment μ

This is the set up of the model. The natural questions to ask now are: does at least one stable matching always exist? When is there a unique stable matching? Can we construct an algorithm to select a stable matching? What are the strategic properties of this algorithm? These questions are addressed in the next section.

3.4 Results

Because this is a model of transferable utility (all agents are risk-neutral with utility functions $u(x) = x$), we know that an equilibrium matching maximizes aggregate output². That is, we know that each partnership produces a joint output. An equilibrium matching is one in which any switching of partners results in a decrease in the sum of output across all pairs.

Recall that the joint output of a partnership (i, j) is:

$$\theta(i, j)Y = \frac{N - \left(\frac{\rho(c_i^j) + \rho(c_j^i)}{2} \right)}{N - 1} Y$$

where $\rho(c_i^j)$ is i 's ranking of j based on i 's evaluation of the compatibility between i 's characteristics and j 's public profile, X_j , and $\rho(c_j^i)$ is j 's ranking of i based on j 's evaluation of the compatibility between j 's characteristics and i 's public profile, X_i .

Thus, it is clear that the matching which maximizes the sum of output across all pairs is also the matching which minimizes the sum of within-partnership rank-sums across all pairs. That is, if i ranks j m^{th} ($\rho(c_i^j) = m$), and j ranks i n^{th} ($\rho(c_j^i) = n$), then the within-partnership rank-sum is $(m + n)$.

This suggests the following approach. The model can be mapped into an alternative model where the utility of an agent i from being matched with an agent j directly depends on i 's ranking of j and j 's ranking of i (where the rankings are based on each agent's evaluation of the public profile of characteristics of potential partners). In particular, let $u_{ij} = -(\rho_i^j + \rho_j^i)$. A profitable block in this model would be if an agent i could instead match with an agent k such that the sum of i 's rank of k and k 's rank of i is *smaller* than the sum of agent i 's rank of j and j 's rank of i . Then, we look for the matching which minimizes the total sum of ranks across all matched pairs. This matching will be the same matching that maximizes aggregate output in the underlying model. The intuition is that in the underlying model, an agent i wants to work with an agent j whom she

²Becker (1973) and Shapley and Shubik (1971) both noted this.

views as compatible, because she knows she will contribute more to a relationship in which she feels satisfied, but she wants agent j to view her as compatible, too—if j doesn't have a high opinion of her, than j will not contribute to the relationship, and the output produced will be low. Hence, i cares about j 's ranking of her as well as her ranking of j .

So, this model can be analyzed by studying the alternative model where each agent forms ordinal preferences over potential partners based off their characteristics, but additionally cares about how those potential partners rank her, so that her utility from a partnership depends equally and directly on her characteristics-based ranking of her partner and on her partner's characteristics-based ranking of her.

To think about existence of equilibrium, and other properties of the equilibrium, it will be helpful to develop several definitions.

Definition 18 The **M matrix** is the $N \times N$ matrix where the $1 \times N$ row vector $M(i, :)$ represents man i 's rankings of women $\{W_1, \dots, W_N\}$. That is, the (i, j) element of the M matrix is the rank that man i assigns to woman j . The **W matrix** is the $N \times N$ matrix where the $N \times 1$ column vector $W(:, j)$ represents woman j 's rankings of men $\{M_1, \dots, M_N\}$. That is, the (i, j) element of the W matrix is the rank that woman j assigns to man i .

Note that if no man finds any woman unacceptable, each row of M will be some permutation of $[1 \dots N]$. Similarly, if no woman finds any man unacceptable, each column of W will be some permutation of $[1 \dots N]$.

While it is the case that in this model, an agent always prefers being matched to somebody rather than remaining single, I suppose that the entity which runs the stable mechanism and to which the ordinal preferences are reported does not know this. Hence an agent is able to misreport a potential partner as being unacceptable by ranking her $2N$.

Definition 19 The **ranks matrix** is the $N \times N$ matrix **ranks** = $M + W$.

In the previous example, the M matrix and the W matrix would be as follows:

$$M : \begin{bmatrix} 3 & 1 & 2 & 4 \\ 1 & 2 & 4 & 3 \\ 4 & 3 & 2 & 1 \\ 1 & 3 & 2 & 4 \end{bmatrix}, \quad W : \begin{bmatrix} 1 & 1 & 2 & 4 \\ 3 & 4 & 1 & 2 \\ 2 & 2 & 3 & 3 \\ 4 & 3 & 4 & 1 \end{bmatrix}$$

Since members of M are along the rows and members of W are along the columns, this tells us that $W2$ is $M3$'s 3rd choice, while $M3$ is $W2$'s 2nd choice.

So, the ranks matrix is:

$$\text{ranks : } \begin{bmatrix} 4 & 2 & 4 & 8 \\ 4 & 6 & 5 & 5 \\ 6 & 5 & 5 & 4 \\ 5 & 6 & 6 & 5 \end{bmatrix}$$

The first result establishes both existence of a stable matching (nonemptiness of the core), and in so doing demonstrates how the DAA can be modified to find a stable matching for this interdependent utility framework.

(A member of M will be referred to as a "man", and a member of W as a "woman", for ease of exposition.)

Proposition 20 (*Existence*) *A stable match with respect to any given set of strict preferences under interdependent utility always exists.*

Proof. The proof is by construction. Consider a modified deferred acceptance algorithm where, in the men-proposing version, men propose first to the woman with whom he would have lowest rank-sum, out of the pool of women he finds acceptable. (If a man finds no women acceptable, he makes no proposals and remains single.) If more than one woman yields the same minimal rank-sum, the man chooses the woman whom he also privately prefers ("selfish tiebreak"). If a woman receives multiple acceptable proposals, she "holds" the man with whom she would have minimal rank-sum, and rejects the other proposals. If two men yield the same minimal rank-sum, she picks the man she also privately prefers. If a woman receives one acceptable proposal, she "holds" that man. If a woman receives only unacceptable proposals, she rejects them. If a woman receives no proposals, she remains single.

In the next round, men who were rejected in the directly previous round propose again, to the woman with whom he would have lowest rank-sum, out of the pool of women who have not rejected him yet and whom he finds acceptable. Women who receive proposals re-evaluate all their options, including men they might have "held" at earlier stages, and keep only the man with whom their rank-sum is smallest (breaking ties selfishly).

This algorithm repeats until no man has been rejected in the directly previous round.

The match this algorithm produces must be stable. Consider a match generated by the algorithm, and suppose that man m has a lower rank-sum with woman w' than with woman w , who is his current partner. Since man m is matched to woman w , he must have found w acceptable (because otherwise he would not have proposed to her), and by transitivity (rationality of preferences) woman w' must also be acceptable to man m . Thus, man m must have proposed to woman w' before proposing to woman w . Since man m is not matched to w' at the end of the algorithm, w' must have rejected man m for a man with whom she has lower rank-sum, or with whom she has the same rank-sum but privately prefers. Thus, woman w' is matched, at the end of the algorithm, to a man she likes more than man m , since preferences are transitive and women cannot do worse as the algorithm progresses, because men propose to them and they hold acceptances. Thus, (m, w') cannot be a blocking pair. (The same logic applies if man m is single at the end of the algorithm, and there exists an acceptable woman w' for him. It must have been that he proposed to w' , but that she rejected him in favor of a man she likes more. Thus, (m, w') cannot be a blocking pair.)

Now suppose that man m has the same rank-sum with woman w' than with woman w , who is his current partner, but m privately prefers w' to w . Since man m is matched to woman w , he must have found w acceptable (because otherwise he would not have proposed to her), and by transitivity (rationality of preferences) woman w' must also be acceptable to man m . Thus, man m must have proposed to woman w' before proposing to woman w , by the selfish tiebreak rule. Since man m is not matched to w' at the end of the algorithm, w' must have rejected man m for a man with whom she has lower rank-sum, or for a man with whom she has the same rank-sum but privately prefers. Thus, woman w' is matched, at the end of the algorithm, to a man she likes more than man m , since preferences are transitive and women cannot do worse as the algorithm progresses, because men propose to them and they hold acceptances. So, (m, w') cannot be a blocking pair.

Since there are no blocking pairs, the match must be stable. ■

This establishes that the core is nonempty in this model, and describes a method for finding a stable matching, given the reported ordinal preferences.

What more can we say about stable matchings in this setting? To think about this, we must partition the set of possible preference profiles.

Definition 21 Given a $T \times T$ matrix A , a matrix B is the vertical reflection of A iff $b_{ij} = a_{i(T-j)}$ for every row i and column j , $i, j \in \{1, \dots, T\}$.

Definition 22 A preference profile $[M \ W]$ is perfectly antagonistic iff M is symmetric ($M = M'$), W is symmetric ($W = W'$), and W is the vertical reflection of M . In other words, for every agent i , $[agent\ i's\ m^{th}\ choice\ ranks\ agent\ i\ n^{th}] \Rightarrow [agent\ i's\ n^{th}\ choice\ ranks\ agent\ i\ m^{th}]$, $m, n \in \{1, \dots, N\}$, $m+n = (N+1)$. Note that in this case, every element of the ranks matrix is equal to $(N+1)$. A preference profile is imperfectly antagonistic iff it is not perfectly antagonistic.

Definition 23 A special case of imperfectly antagonistic preferences is mutual x preferences, where, given some $x \in \{1, \dots, N\}$, every agent's x^{th} choice ranks him/her x^{th} . An agent's preferences are mutual first if his/her first choice ranks him/her first.

(The case where every agent has mutual first preferences is not interesting, since there is no within-side or between-side conflict of any kind, and the optimal matching is obvious—everybody gets his or her first choice.)

There exist also preference profiles where a subset of agents have perfectly antagonistic preferences, over all agents in that subset.

Definition 24 A preference profile is essentially perfectly antagonistic if players' preferences can be partitioned into mutual first preferences and perfectly antagonistic preferences. That is, $M_f \subset M$, $W_f \subset W$, $|M_f| = |W_f| = K \in \{1, 2, \dots, N\}$, where M_f, W_f have mutual first preferences for each other, while $M \setminus M_f$ and $W \setminus W_f$, both of cardinality $(N - K)$, have perfectly antagonistic preferences for each other. A preference profile is essentially imperfectly antagonistic iff it is not essentially perfectly antagonistic.

This partition of preference profiles is specific to the structure of interdependence in this model. Because the model maps into a framework where an agent's utility from a match depends directly and equally on each partner's rank of the other, and in particular the dependence is linear, we must treat specially sets of preference profiles where the rank-sum for all possible partnerships within a weak subset of agents in M and W is exactly the same.

For example, the classic case of perfect antagonism is the following:

$$M1 : W1 > W2$$

$$M2 : W2 > W1$$

$$W1 : M2 > M1$$

$$W2 : M1 > M2$$

$M1$'s favorite woman is $W1$, but her favorite man is $M2$, whose favorite woman is $W2$, whose favorite man is $M1$. Thus, the ranks matrix is:

$$ranks : \begin{bmatrix} 3 & 3 \\ 3 & 3 \end{bmatrix}$$

and both of the possible matches are stable.

Note that the antagonism is "perfect", because not only does everyone's first choice prefer someone else, but if the match were $(M1, W1)$ and $(M2, W2)$, then making $M1$ happier by partnering him with $W2$ decreases $W2$'s happiness by exactly the amount that $M1$'s happiness was increased. That is, no one's happiness can be increased via a change of partner without decreasing someone else's happiness by an exactly-offsetting amount.

The next result shows that there is a unique stable matching given essentially imperfectly antagonistic preference profiles, while there are multiple stable matchings given essentially perfectly antagonistic preference profiles.

Proposition 25 (a) *(Essentially imperfectly antagonistic preferences) The deferred acceptance algorithm generates the same match when men propose as when women propose. Further, the generated match is the unique stable match with respect to the preference profile, under interdependent utility.*

(b) *(Essentially perfectly antagonistic preferences) There are multiple stable matchings: the K men and K women who mutually prefer each other first are always matched with each other, and the $(N - K)$ men and women with perfectly antagonistic preferences are matched with each other such that all men receive their m^{th} choice, and all women receive their n^{th} choice, where $m \in \{1, \dots, N - K\}, n \in \{1, \dots, N - K\}, m + n = N - K + 1$.*

Proof. (a) Consider any given imperfectly antagonistic preference profile. Let a stable matching with respect to this preference profile be denoted μ . By definition of stability, we know that every

agent is matched to an acceptable partner, and there are no blocking pairs: that is, there is no pair (i, j) , $i \in M$, $j \in W$, such that $-(rank_i(\mu(i)) + rank_{\mu(i)}(i)) < -(rank_i(j) + rank_j(i))$ and $-(rank_j(\mu(j)) + rank_{\mu(j)}(j)) < -(rank_i(j) + rank_j(i))$. Now consider a different matching, obtained by switching the partners of two men, i and i' , so that i is now matched with $\mu(i')$ and i' is matched with $\mu(i)$. But then it must be that *both* $(i, \mu(i'))$ and $(i', \mu(i))$ are worse off—if $(i, \mu(i'))$ became better off, they would have blocked the original match, and if $(i', \mu(i))$ became better off, they would have blocked the original match. But then this new match cannot be stable, since $(i, \mu(i))$ and $(i', \mu(i'))$ are blocking pairs. Since any possible matching can be achieved by switching two partners in the original match, it follows the the original stable matching μ must be the unique stable match.

Since it was proven in the previous proposition that the deferred acceptance algorithm (regardless of which side proposes) generates a stable matching, it must be that the deferred acceptance algorithm generates the same match when men propose as when women propose.

(b) It is clear that in any stable matching, mutual firsts are matched to each other. Consider the remaining $(N - K)$ agents, who have perfectly antagonistic preferences over each other, and let their section of the M , W , and $rank$ s matrices be denoted M_T , W_T , and $rank_{sT}$, respectively. Then, by definition, $M_T = M'_T$, $W_T = W'_T$, M_T is the vertical reflection of W_T , and every element of $rank_{sT}$ is the same, and equal to $(N - K + 1)$. Hence, every agent gets the same base utility $-u(x, y) = (N - K + 1)$ from being matched with any partner. Moreover, in any possible matching, all men are matched to their m^{th} choice, while all women are matched to their $(N - K + 1 - m)^{th}$ choice. The selfish tiebreak rule means that man i and woman j can be a blocking pair iff man i ranks j more highly than his given partner, *and* woman j ranks man i more highly than her given partner. But, by definition of perfectly antagonistic preferences, every agent on one side's gain comes at their partner on the other side's loss. Thus, there can be no blocking pairs from matches where mutual first agents are matched to each other, and among the remaining perfectly antagonistic agents, each man is matched to his m^{th} choice, and each woman is matched to her $(N - K + 1 - m)^{th}$ choice. ■

Hence, many preference profiles which yield multiple stable matchings in the standard framework (some of which are better for M and some of which are better for W), and which therefore seem characterized by a seemingly irreconcilable level of conflict, may not be so irreconcilable after all, if we account for interdependence of utility. Except in cases of perfect antagonism (where an

increase in anybody's happiness is exactly offset by a decrease in someone else's happiness), there is a unique stable matching, and therefore it is irrelevant which side proposes. Moreover, although there are multiple stable matchings in cases of perfect antagonism, market participants are indifferent over those matchings—it isn't the case that some of the matchings are better for one set of people, while others are better for another set of people.

Not only is the stable matching unique given essentially imperfectly antagonistic preference profiles, but it has a special property—it is either the man- or the woman-optimal stable matching from the standard framework.

Proposition 26 *The unique stable match with respect to a given (essentially) imperfectly antagonistic preference profile under interdependent utility is either the man- or the woman-optimal stable match with respect to preferences under standard utility.*

Proof. We know that, in the standard model, the woman-optimal match gives every woman her best achievable mate and the man his worst achievable mate; the man-optimal match gives every man his best achievable mate and every woman her worst achievable mate. "Achievable" here is in the standard sense of "most preferred mate among the complete set of stable matches" (where stability is based purely on independent utility, so that a man and woman block iff they each rank each other higher than their original partners). Clearly, in the independent utility framework, every woman's weakly favorite match is the woman-optimal match, and every man's weakly favorite match is the man-optimal match, where these two matches generally differ. However, if agents have linearly interdependent utility, the woman-optimal match is the match where each woman is matched to her best "achievable" mate, and the man is matched to a woman *who likes him to the degree* that he is her *best* "achievable" mate. Similarly, in the man-optimal match, the man is matched to his best "achievable" mate, and the woman is matched to a man *who likes her to the degree* that she is his *best* "achievable" mate. Thus, all other matches that are not the man- or the woman-optimal match must do strictly worse under interdependent utility, since men and women not being matched to their best achievable mates implies that men and women are also not matched to someone who desires them to the degree that they are their best achievable mate. And, as long as preferences are not (essentially) perfectly antagonistic, either the man- or the woman-optimal must do strictly better. Therefore, the unique stable match with respect to a given imperfectly antagonistic preference profile under interdependent utility must be either the man- or woman-optimal stable match with respect to the same preferences, under independent utility. ■

Finally, what can we say about strategic properties of the stable mechanism? We know in the standard case that the core is nonempty, and that the set of stable matchings is generally not unique—in particular, the man-optimal stable matching differs from the woman-optimal stable matching. We know that under a stable mechanism, truthtelling is never a dominant strategy, and moreover, if there are multiple stable matchings, then at least one person has an incentive to misreport preferences, assuming the other are telling the truth.

By contrast, we know that if agents' utilities are interdependent, in the sense that they care about the contributions made by each partner to the relationship, where a partner's contributions are increasing in satisfaction with the relationship, then the core is nonempty, and the stable matching is generally unique. Multiple stable matchings arise only in cases of perfectly antagonistic preference profiles, and in that case, all agents are simply indifferent over those matchings.

The natural question to ask is, will at least one agent have an incentive to misreport preferences under a stable mechanism in this model, assuming all other participants are telling the truth? The answer turns out to be no.

Proposition 27 *In the basic model with interdependent utility and (essentially) imperfectly antagonistic preferences, as long as no agent is pivotally antagonistic, no individual agent has an incentive to misreport preferences if all other agents are telling the truth. Therefore, given a stable rule, it is always an equilibrium for an agent to state his or her true preferences.*

Proof. Suppose that the preference profile is imperfectly antagonistic. Suppose an agent misreports his preferences while all other agents report the truth. The principal uses a stable matching mechanism, so the resulting match will be the *unique* stable match with respect to the *false* preferences, where the agent is matched to a different partner than he would have been had he reported truthfully (else, there would have been no reason to lie). But by Proposition 9(a), we know that this agent and the partner he would have had had he reported truthfully must be a blocking pair in the unique stable match with respect to the false preferences. That is, the agent is better off with his truthful partner than the partner he gets by lying. Hence, truthtelling is an equilibrium strategy.

Now, suppose the full preference profile is essentially perfectly antagonistic. Then all agents are indifferent over the multiple stable matchings, so no agent has an incentive to lie. ■

Hence, if agents' utilities are interdependent in the sense of this model, then truthful reporting of preferences is a best response under a stable mechanism such as a modified DAA, if everyone else

is reporting truthfully. The across-group disagreement that resulted in multiple stable matchings in the standard framework is resolved by interdependence—an agent no longer cares only about being matched with someone as high up her list as possible. She also values her partner’s opinion of her, since that is her only indication of the contribution her partner would make to a joint relationship, and if her partner is dissatisfied, the contribution, and subsequently the output, will be low.

Thus, our observation that stable mechanisms work well in practice despite the seemingly various ways to manipulate profitably in theory might be due to the fact that agents care about their desirability as well as their desires. If this is the case, then I’ve shown that there are no ways to gain by misreporting preferences under a stable mechanism, including reporting truncated preferences. So, one interpretation of these results is that unaccounted-for interdependence explains the smooth functioning of stable mechanisms.

Another interpretation is that, if we think there *is* interdependence of agents’ utilities, (or we know there is interdependence based on empirical or experimental work), then we should implement the stable mechanism described in the existence proof. That is, we should ask both sides to report their ordinal preferences, and then run the modified DAA on the ranks matrix to find the unique stable matching (which side proposes is irrelevant). We can honestly assure market participants that truthtelling is an optimal strategy.

3.5 An Example: Followup

Recall the ordinal preferences from the first example:

$$M1 : W2 > W3 > W1 > W4$$

$$M2 : W1 > W2 > W4 > W3$$

$$M3 : W4 > W3 > W2 > W1$$

$$M4 : W1 > W3 > W2 > W4$$

$$W1 : M1 > M3 > M2 > M4$$

$$W2 : M1 > M3 > M4 > M2$$

$$W3 : M2 > M1 > M3 > M4$$

$$W4 : M4 > M2 > M3 > M1$$

I showed that the man-optimal and the woman-optimal stable matching differed:

$$\mu_M = \begin{bmatrix} M1 & M2 & M3 & M4 \\ W2 & W1 & W4 & W3 \end{bmatrix}$$

$$\mu_W = \begin{bmatrix} M1 & M2 & M3 & M4 \\ W2 & W3 & W1 & W4 \end{bmatrix}$$

Thus, at least one person has an incentive to misreport preferences under a stable mechanism. For example, if the man-optimal DAA is the stable mechanism, then $W4$ can gain by reporting only $M4$ to be acceptable to her.

Now suppose that agents have interdependent utility as described by this model. Then the ranks matrix is:

$$ranks : \begin{bmatrix} 4 & 2 & 4 & 8 \\ 4 & 6 & 5 & 5 \\ 6 & 5 & 5 & 4 \\ 5 & 6 & 6 & 5 \end{bmatrix}$$

Suppose men propose. Then the first round is:

$$M1 \rightarrow W2$$

$$M2 \rightarrow W1$$

$$M3 \rightarrow W4$$

$$M4 \rightarrow W1$$

$W1$ holds $M2$'s proposal and rejects $M4$'s, so he proposes to his next choice, $W4$, and the next round is:

$$M1 \rightarrow W2$$

$$M2 \rightarrow W1$$

$$M3 \rightarrow W4$$

$$M4 \rightarrow W4$$

$W4$ holds $M3$'s proposal and rejects $M4$'s, so he proposes to $W3$, and the stable matching is:

$$\begin{bmatrix} M1 & M2 & M3 & M4 \\ W2 & W1 & W4 & W3 \end{bmatrix}$$

And when women propose, the algorithm finds the same stable match. This is the unique stable matching.

Note that this match is the man-optimal stable match from the standard framework.

Now, let's explore strategic aspects. For example, does $W4$ still have an incentive to report $M4$ as being her only acceptable partner?

Suppose she ranks $M4$, 1, and the other three men, 8 (recall that the protocol for declaring someone unacceptable is to rank him $2N$).

Then the ranks matrix is:

$$ranks : \begin{bmatrix} 4 & 2 & 4 & 12 \\ 4 & 6 & 5 & 11 \\ 6 & 5 & 5 & 9 \\ 5 & 6 & 6 & 5 \end{bmatrix}$$

and the match found by the modified DAA is:

$$\begin{bmatrix} M1 & M2 & M3 & M4 \\ W2 & W1 & W3 & W4 \end{bmatrix}$$

But this lie is not profitable for $W4$, precisely because of interdependent utility—while $W4$ does rank $M4$ more highly than $M3$, her previous partner, $M4$ ranks her last. By contrast, $M3$ ranks

her first, and she ranks him third. If $W4$ were to partner with $M4$, the output they would produce would be $\frac{1}{2}Y$, while if $W4$ partnered with $M3$, the output they would produce would be $\frac{2}{3}Y$. Hence, $W4$ actually prefers $M3$ as a partner over $M4$ — $M3$ would contribute highly to their relationship, while $M4$ wouldn't contribute at all.

3.6 Conclusion

Dubins and Freedman (1981) and Roth (1982) showed that it's impossible to design a stable mechanism under which truth-telling is a dominant strategy for every agent, or even a best response for every agent, assuming the others are telling the truth. Hence, if we wish to continue using stable mechanisms, research must focus on understanding why market participants do not misreport their preferences, even when they have opportunities to gain by doing so.

In this paper, I offer an explanation which departs from the existing work on computational complexity, incomplete information, and large markets. I show that the utility of market participants may in fact be interdependent—an agent may care about a potential partner's ranking of her, in addition to her ranking of her partner. I show that when the interdependence is linear and desires and desirability are equally-weighted, there is either a unique stable matching, or multiple stable matchings across which all agents are indifferent. Thus, no agent has an incentive to lie under a stable mechanism, if the other participants are telling the truth.

There are several important caveats to these results, and some important directions for future research. In this paper, interdependence is modeled in a very specific way, and this leads to a partition of preference profiles which is specific to the structure of interdependence (that is, perfect antagonism is a concept which seems specific to this model). However, the intuition for uniqueness does not seem to be bound by the functional form. It would be of great interest to generalize the concept of interdependent utility.

Experimental work would benefit this research greatly. There is no stable mechanism under which truth-telling is a best response to others telling the truth in the standard model, but I've shown that there is such a truth-inducing stable mechanism if agents care sufficiently about their partners' perceptions. It would therefore be of interest to run the following simple one-to-one, two-sided matching experiment: in the first round, each participant in the experiment is shown the profiles of characteristics of all the members in the other group, and asked to submit their rankings of the members in the other group. Then, it is revealed to each participant i how she was ranked

by each member of the other group. Agents would then be allowed to re-submit their preference lists. Are there differences between the original and the updated preference lists? Do the differences appear to have a special structure?

Currently, our way of dealing with the possibility of strategic manipulation is to push two-sided problems to be somewhat one-sided. In the NRMP, only residents submit rankings over hospitals, while in student-school matching problems, only students are allowed to rank schools. (Hospitals and schools are still allowed to set quotas.) This somewhat suppresses the two-sidedness of the problem, and induces residents and students to report truthfully. However, if our hope is for market design to expand and find new applications, then we need to have a better understanding of strategic manipulation, and better ways of dealing with it. This paper suggests that, contrary to existing explanations where agents have an incentive to misreport but cannot bear the cost of calculating the optimal lie, or do not have enough information, or gain little from it, it may be the case that agents do not have an incentive to misreport after all. Agents know that the partners they could get by lying are precisely those that they desire but who do not desire them, and they may recognize that such a partnership might not be so satisfying. This paper identifies one structure of interdependence in which the stable matching is unique, and constructs a mechanism to find that stable matching. Given how powerful interdependence can be, a fruitful next step would be to try and experimentally identify the structure of interdependence in real-life matching problems, and then to incorporate these observations into further theoretical research.

Bibliography

- [1] "2012 National Trade Estimate Report: Foreign Trade Barriers." U.S. Wheat Associates Oct. 2011.
- [2] Akerberg, D. and M. Botticini. "Endogenous Matching and the Empirical Determinants of Contract Form." *Journal of Political Economy* (110) 2002 564-590.
- [3] Akanda, M., H. Isoda, and S. Ito. "Problem of Sharecrop Tenancy System in Rice Farming in Bangladesh: A Case Study on Alinapara Village in Sherpur District." *Journal of Int. Farm Management* (4) 2008 23-30.
- [4] Alderman, H. and C. Paxson. "Do the Poor Insure? A Synthesis of the Literature on Risk and Consumption in Developing Countries." *Economics in a Changing World*, vol. 4 1992 49-78.
- [5] Allen, D. and D. Lueck. "Contract Choice in Modern Agriculture: Cash Rent versus Crop-share." *Journal of Law and Economics* (35) 1992 397-426.
- [6] Ambrus, A., M. Mobius, and A. Szeidl. "Consumption Risk-sharing in Social Networks." *American Economic Review* forthcoming.
- [7] Antle, J. "Econometric Estimation of Producers' Risk Attitudes." *Amer. Journal of Ag. Econ.* (69) 1987 509-522.
- [8] Attanasio, O., A. Barr, J. Cardenas, G. Genicot, and C. Meghir. "Risk Pooling, Risk Preferences, and Social Networks." *Amer. Econ. Journal: Applied Economics* (4) 2012 132-167.
- [9] Attanasio, O. and J. Rios-Rull. "Consumption Smoothing in Island Economies: Can Public Insurance Reduce Welfare?" *European Economic Review* (44) 2000 1225-1258.
- [10] Banerjee, A. "Prospects and Strategies for Land Reforms." *Annual World Bank Conference on Development Economics 1999*, eds. B. Pleskovic and J. Stiglitz. 253-284.

- [11] Banerjee, A., P. Gertler, and M. Ghatak. "Empowerment and Efficiency: Tenancy Reform in West Bengal." *Journal of Political Economy* (110) 2002 239-280.
- [12] Bardhan, P. "Interlocking Factor Markets and Agrarian Development: A Review of Issues." *Oxford Econ. Papers* (32) 1980 82-98.
- [13] Basu, K. "Limited Liability and the Existence of Share Tenancy." *J. Dev. Econ.* (38) 1992 203-220.
- [14] Becker, G.S. "A Theory of Marriage: Part I." *Journal of Political Economy* (81) 1973 813-846.
- [15] Bellemare, M. "Sharecropping, Insecure Land Rights and Land Titling Policies: A Case Study of Lac Alaotra, Madagascar." *Development Policy Review* (27) 2009 87-106.
- [16] Besley, T. "Savings, Credit, and Insurance." *Handbook of Development Economics* (3) 1995 2123-2207.
- [17] Binswanger, H.P. "Attitudes towards Risk: Experimental Measurement in Rural India." *Am. J. Agr. Econ.* (62) 1980 395-407.
- [18] Bloch, F., G. Genicot, and D. Ray. "Informal Insurance in Social Networks." *J. Econ. Theory* (14) 2008 36-58.
- [19] Bold, T. "Implications of Endogenous Group Formation for Efficient Risk-Sharing." *The Economic Journal* (119) 2009 562-591.
- [20] Bramoulle, Y. and R. Kranton. "Risk Sharing Networks." *J. Econ. Behavior and Organization* (64) 2007 275-294.
- [21] Chakraborty, A., A. Citanna, and M. Ostrovsky. "Two-sided matching with interdependent values." *Journal of Economic Theory* (145) 2010 85-105.
- [22] Chambers, C. and F. Echenique. "When Does Aggregation Reduce Risk Aversion?" *Games and Economic Behavior* (76) 2012 582-595.
- [23] Chandrasekhar, A. et al. "Testing Models of Social Learning on Networks: Evidence from a Framed Field Experiment." *Working Paper* 2012.
- [24] Chiappori, P-A., K. Samphantharak, S. Schulhofer-Wohl, and R. Townsend. "Heterogeneity and Risk-Sharing in Village Economies." *Working Paper* 2011.
- [25] Chiappori, P-A. and P. Reny. "Matching to share Risk." *Working Paper* 2006.

- [26] Chung, K.S. "Implementing stable matchings under incomplete information." PhD Dissertation, Chapter 2 1999.
- [27] Dalla Marta, A., M. Mancini, R. Ferrise, M. Bindi, and S. Orlandini. "Energy crops for biofuel production: Analysis of the potential in Tuscany." *Biomass and Bioenergy* (34) 2010 1041-1052.
- [28] Dawe, D. "How far down the path to free trade? The importance of rice price stabilization in developing Asia." *Food Policy* (26) 2001 163-175.
- [29] Dercon, S. "Risk, Crop Choice, and Savings: Evidence from Tanzania." *Econ. Dev. and Cultural Change* (44) 1996 483-513.
- [30] Dercon, S. and P. Krishnan. "In Sickness and in Health: Risk Sharing within Households in Rural Ethiopia." *Journal of Political Economy* (108) 2000 688-727.
- [31] Dubins, L. and D. Freedman. "Machiavelli and the Gale-Shapley algorithm." *American Mathematical Monthly* (88) 1981 485-494.
- [32] Eeckhout, J. and K. Munshi. "Matching in Informal Financial Institutions." *Journal of the European Economic Association* (8) 2010 947-988.
- [33] Fafchamps, M. "Risk-sharing Between Households." *Handbook of Social Economics* 2008.
- [34] Fafchamps, M. and A. Lund. "Risk-sharing Networks in Rural Phillippines." *J. Dev. Econ.* (71) 2003 261-287.
- [35] Fafchamps, M. and F. Gubert. "The Formation of Risk-Sharing Networks." *J. Dev. Econ.* (83) 2007 326-350.
- [36] Franco, A., M. Mitchell, and G. Vereshchagina. "Incentives and the Structure of Teams." *Journal of Economic Theory* (146) 2011 2307-2332.
- [37] Gale, D. and L. Shapley. "College admissions and the stability of marriage." *American Mathematical Monthly* (69) 1962 9-15.
- [38] Gale, D. and M. Sotomayor. "Ms. Machiavelli and the stable matching problem." *American Mathematical Monthly* (92) 1985 261-268.
- [39] Genicot, G. and D. Ray. "Group Formation in Risk-Sharing Arrangements." *Rev. Econ. Stud* (70) 2003 87-113.

- [40] Grinstead, C. and J. Snell. "Introduction to Probability." American Mathematical Society 1997.
- [41] Holmstrom, B. and P. Milgrom. "Aggregation and Linearity." *Econometrica* (55) 1987 303-328.
- [42] Jewitt, I. "Justifying the First-Order Approach to Principal-Agent Problems." *Econometrica* (56) 1988 1177-1190.
- [43] Jewitt, I. "Risk Aversion and the Choice Between Risky Prospects: the Preservation of Comparative Statics Results." *Review of Economic Studies* (54) 1987 73-85.
- [44] *Insurance Against Poverty*. ed. Stefan Dercon. Oxford University Press. 2005.
- [45] Kaneko, M. "The Central Assignment Game and the Assignment Markets." *Journal of Mathematical Economics* (10) 1982 205-232.
- [46] Kara, T. and T. Sonmez. "Nash implementation of matching rules." *Journal of Economic Theory* (68) 1996 425-439.
- [47] Knudsen, O. and J. Nash. "Domestic Price Stabilization Schemes in Developing Countries." *Economic Development and Cultural Change* (38) 1990 539-558.
- [48] Kocherlakota, N. "Implications of Efficient Risk-Sharing Without Commitment." *Review of Economic Studies* (63) 1996 596-609.
- [49] Kojima, F. and P. Pathak. "Incentives and stability in large two-sided matching markets." *American Economic Review* (99) 2009 608-627.
- [50] Kurosaki, T. "Risk-Sharing Arrangements and the Structure of Risk and Time Preferences: Theory and Evidence from Village India." Working Paper 1999.
- [51] Legros, P. and A. Newman. "Beauty is a Beast, Frog is a Prince: Assortative Matching with Nontransferabilities (75) 2007 1073-1102.
- [52] Li, Sunxi, H. Sun, and P. Chen. "Assortative Matching of Risk-Averse Agents with Endogenous Risk." *Journal of Economics* (online only) 2012.
- [53] Ligon, E., J. Thomas, and T. Worrall. "Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies." *Review of Economic Studies* (69) 2002 209-244.
- [54] Mazzocco, M. "Household intertemporal behavior: a collective characterization and a test of commitment." *Review of Economic Studies* (74) 2007 857-895.

- [55] Minot, N. "Food Price Stabilization: Lessons for Eastern and Southern Africa." Risk Management in African Agriculture Seminar Presentation 2010.
- [56] Mirrlees, J. "The Implications of Moral Hazard for Optimal Insurance." Mimeo 1979.
- [57] Morduch, J. "Income Smoothing and Consumption Smoothing." J. Econ. Perspectives (9) 1995 103-114.
- [58] Mumcu, A. and I. Saglam. "One-to-one matching with interdependent preferences." MPRA Paper No. 1908 2007.
- [59] Pini, M., F. Rossi, K. Venable, and T. Walsh. "Manipulation and gender neutrality in stable marriage procedures." AAMAS '09: Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems 2009 655-672.
- [60] Rabin, M. "Risk aversion and expected utility theory: A calibration theorem." Econometrica (68) 2000 1281-1292.
- [61] Rathore, V.S., J.P. Singh, and R.K. Beniwal. "Rainfed Farming Systems of Hyper Arid Northwestern Rajasthan: an Analysis." Annals of the Arid Zone (48) 2009 125-131.
- [62] Rogerson, W. "The First-Order Approach to Principal-Agent Problems." Econometrica (53) 1985 1357-1367.
- [63] Rosenzweig, M. "Risk, Implicit Contracts, and the Family in Rural Areas of Low-Income Countries." Economic Journal (98) 1988 1148-1170.
- [64] Rosenzweig, M. and H.P. Binswanger. "Wealth, Weather Risk, and the Composition and Profitability of Agricultural Investments." The Economic Journal (103) 1993 56-78.
- [65] Rosenzweig, M. and O. Stark. "Consumption Smoothing, Migration, and Marriage: Evidence from Rural India." J. Polit. Econ. (97) 1989 905-926.
- [66] Roth, A. "The economics of matching stability and incentives." Mathematics of Operations Research (7) 1982 617-628.
- [67] Roth, A. "Misrepresentation and stability in the marriage problem." Journal of Economic Theory (34) 1984 383-387.
- [68] Roth, A. "Matching with incomplete information." Games and Economic Behavior (1) 1989 191-209.

- [69] Roth, A. "The economist as engineer: Game theory, experimentation, and computation as tools for design economics." *Econometrica* (70) 2002 1341-78.
- [70] Roth, A. and U. Rothblum. "Truncation strategies in matching markets—in search of advice for participants." *Econometrica* (67) 1999 21-43.
- [71] Roth, A. and M. Sotomayor. "Two-sided matching." Cambridge University Press, Cambridge, MA, 1990.
- [72] Schulhofer-Wohl, S. "Negative-assortative Matching of Risk-Averse Agents with Transferable Expected Utility." *Econ. Letters* (92) 2006 383-388.
- [73] Serfes, K. "Endogenous Matching in a Market with Heterogeneous Principals and Agents." *Int J Game Theory* (36) 2008 587-619.
- [74] Shapley, L. and M. Shubik. "The assignment game I: The Core." *Intl. J. Game Theory* (1) 1971 111-130.
- [75] Sharpe, William F. "The Sharpe Ratio". *The Journal of Portfolio Management* (21) 1994 49-58.
- [76] Singh, H., G. Singh, J. Samuel, and R. Pathak. "Precision Farming in Horticulture." *Proceedings of the National Seminar Workshop on Hi-Tech Horticulture and Precision Farming*, 2003.
- [77] Stiglitz 1974. "Incentives and Risk Sharing in Sharecropping." *Review of Economic Studies* (41) 1974 219-255.
- [78] "The Origins of Tuscan Landscape." *The Cultured Traveler* Jan. 2006.
- [79] Townsend, R. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory* (21) 1979 265-293.
- [80] Townsend, R. "Risk and Insurance in Village India." *Econometrica* (62) 1994 539-591.
- [81] Townsend, R. and R. Mueller. "Mechanism Design and Village Economies: From Credit to Tenancy to Cropping Groups." *Review of Economic Dynamics* (1) 1998 119-172.
- [82] Wang, X. "Endogenous Insurance and Informal Relationships." Working Paper 2012a.
- [83] Wang, X. "A Note on Moral Hazard and Linear Compensation Schemes." Working Paper 2012b.
- [84] Wang, X. "Risk, Incentives, and Contracting Relationships." Working Paper 2012c.

- [85] *Rainfed Agriculture: Unlocking the Potential*. ed. S. Wani, J. Rockstrom, and T. Oweis. CAB International. 2009.
- [86] Wilson, R. "The Theory of Syndicates." *Econometrica* (36) 1968 119-132.
- [87] Wright, B. "Storage and Price Stabilization." *Handbook of Ag. Econ.* (1) 2001 818-861.